

A;Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-501 <WUX>
A:Cross-references: EMBL:U55376; NID:g1280130; PID:g1280135; PIDN:AAA98007.1; GSPDB:GN01
A:Experimental source: strain Bristol N2; clone F16H11
C:Genetics:
A:Gene: CBSP:F16H11.1
A:Map position: X
A:Introns: 18/3; 41/1; 81/3; 117/1; 163/3; 187/2; 224/3; 255/2; 283/3; 314/3; 346/3; 40:

Query Match 39.7%; Score 48; DB 2; Length 501;
Best Local Similarity 45.0%; Pred. No. 13;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 LQLARCYLSNSOAYSAYYI 22
| : | : | : | : | : ||
Db 287 LYMSRLYNISQVYPFYI 306
| : | : | : | : | : ||

RESULT 5
C97183
probable nucleotidyltransferase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: C97183
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
J.; Daly, M.J.; Bennett, G.N.; Kennin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A:Reference number: A96900; PMID:21359325; PMID:21359325
A:Accession: C97183
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-298 <GUR>
A:Cross-references: GB:AE001437; PIDN:AAK80254.1; PID:gl5025303; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2298

Query Match 38.8%; Score 47; DB 2; Length 298;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 6 LAECYLNSOAYSAYY 21
| : | : | : | : | : ||
Db. 142 LCKKYLNNKKRYTAY 157
| : | : | : | : | : ||

RESULT 6
C71066
hypothetical protein PH1225 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Nov-1999
C:Accession: C71066
R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch
DNA Res. 5, 55-76, 1998
A:title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; PMID:98344137; PMID:9679194
A:Accession: C71066
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-173 <RAW>
A:Cross-references: GB:AP000005; NID:g3236132; PIDN:BAA30325.1; PID:d1031268; PID:g32327
A:Experimental source: strain OT3
A>Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PHI225

Query Match 38.0%; Score 46; DB 2; Length 173;
Best Local Similarity 45.8%; Pred. No. 9.6;
Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 5 LLARCYLNSQAYSAY 20
 :|||||: :
 Db 936 MLARCYVRNKKPQSAW 951

RESULT 12

F64038
 Hypothetical protein H11622 - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
 C:Accession: F64038
 R:Feischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
 ; Gocayne, J.D.; Scott, J.; Shirley, J.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
 ; D.M.; Branton, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Gogghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: F64038
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-161 <TIGR>
 A:Cross-references: GB:U32835; GB:I42023; NID:G1574459; PIDN:AA23271.1; PID:G1574471;
 C:Superfamily: Haemophilus influenzae hypothetical protein H11622

Query Match 37.2%; Score 45; DB 2; Length 161;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 LLARCYLNSQAYSAY 20
 :|||||: :
 Db 10 LFAACYLPNAVHAY 25

RESULT 13

S66765
 Hypothetical protein YOL072w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein O1140
 C:Species: Saccharomyces cerevisiae
 C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
 C:Accession: S66765
 R:Alexandraki, D.; Katsoulou, C.; Tzermia, M.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66756
 A:Accession: S66765
 A:Molecule type: DNA
 A:Residues: 1-455 <ALE>
 A:Cross-references: EMBL:Z74814; NID:G1419898; PID:G1419899; GSPDB:GN00015; MIPS:YOL072w
 A:Experimental source: strain S288C

C:Genetics:
 A:Gene: SGD:THP1; MIPS:YOL072w
 A:Cross-references: SGD:S0005433
 A:Map position: 15L
 C:Superfamily: Saccharomyces hypothetical protein YOL072w

Query Match 37.2%; Score 45; DB 2; Length 455;
 Best Local Similarity 56.2%; Pred. No. 35;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 LLARCYLNSQAYSAY 20
 :|||||: :
 Db 225 LLGRVYLLNSQVHNAF 240

RESULT 14

T24088
 Hypothetical protein R09E10.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T24088
 R:Matthews, L.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z19839

A:Accession: T24088
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1459 <WIL>
 A:Cross-references: EMBL:Z70287; PIDN:CAA94300.1; GSPDB:GN00022; CESP:R09E10.5
 A:Experimental source: Clone R09E10
 C:Genetics:
 A:Gene: CESP:R09E10.5
 A:Map position: 4
 A:Introns: 86/1; 128/3; 392/1; 504/3; 723/3; 763/2; 819/3; 865/1; 1108/1; 1168/2; 1197/

Query Match 37.2%; Score 45; DB 2; Length 1459;
 Best Local Similarity 58.3%; Pred. No. 1.1e-02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 CYLSNSQAYSAY 20
 :|||||: :
 Db 739 CYMSSNTYSSY 750

RESULT 15

S39835
 Hypothetical protein YBL052c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YBL0507; hypothetical protein YBL0515
 C:Species: Saccharomyces cerevisiae
 C:Date: 16-Dec-1993 #sequence_revision 09-Sep-1994 #text_change 29-Oct-1999
 C:Accession: S39835; S45787; S37336
 R:Scherens, B.; el Bakkoury, M.; Vierendeels, F.; Dubois, E.; Messenguy, F.
 Yeast 9, 1355-1371, 1993
 A:Title: Sequencing and functional analysis of a 32 560 bp segment on the left arm of
 A:Reference number: S39824; MUID:94205266; PMID:8154187
 A:Accession: S39835
 A:Molecule type: DNA
 A:Residues: 1-831 <SCH>
 A:Cross-references: EMBL:Z23261; NID:G313733; PIDN:CAA80794.1; PID:G313745
 A:Experimental source: strain S288C
 R:Dubois, E.; el Bakkoury, M.; Glansdorff, N.; Messenguy, F.; Pierard, A.; Scherens, B.
 submitted to the Protein Sequence Database, August 1994

A:Reference number: S45782
 A:Accession: S45787
 A:Molecule type: DNA
 A:Residues: 1-831 <DUB>
 A:Cross-references: EMBL:Z35814; NID:G536078; PIDN:CAA84873.1; PID:G536080; MIPS:YBL0505
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:SAS3
 A:Cross-references: SGD:S0000148; MIPS:YBL052C
 A:Map position: 2L

Query Match 36.8%; Score 44.5; DB 2; Length 831;
 Best Local Similarity 37.0%; Pred. No. 74;
 Matches 10; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

QY 2 NLQLLARCYLNSQAY-----SAYVIL 23
 :|||||: :
 Db 355 NLCLLACPFINSKTLVDYVEPFVIL 381

Search completed: February 2, 2004, 13:02:38
 Job time : 69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 2, 2004, 08:49:58 ; Search time 38 Seconds
(without alignments)
29.701 Million cell updates/sec

Title: US-10-036-492-6

Perfect score: 121

Sequence: 1 VNLQLARCYLSNQAYSAYILK 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	48.8	824	1 CC27 HUMAN	P30260 homo sapien
2	46	38.0	806	1 BIMA EMENI	P17885 emericeella
3	45	37.2	161	1 YG22 HAEIN	P44275 haemophilus
4	45	37.2	1459	1 YF1M CABEL	Q21874 caenorhabdi
5	44.5	36.8	831	1 SAS3 YEAST	P34218 saccharomyc
6	44	36.4	504	1 PSD3 CABEL	Q04908 caenorhabdi
7	43.5	36.0	1102	1 YG49 SCHPO	Q80184 schizosacch
8	43	35.5	266	1 Y302 STRAW	Q8nyb7 staphylococ
9	43	35.5	266	1 Y325 STAAH	Q99wq1 staphylococ
10	43	35.5	325	1 MTE1 ECOLI	P00472 escherichia
11	43	35.5	441	1 COAT SOGMV	P15627 soybean chl
12	43	35.5	513	1 TI60 HUMAN	Q32993 homo sapien
13	43	35.5	558	1 VB04 VACCC	P21001 vaccinia vi
14	43	35.5	558	1 VB04 VACCC	P24789 vaccinia vi
15	43	35.5	932	1 SYL ARCFU	Q30250 archaeoglob
16	42	34.7	147	1 NDD BFR70	P42268 bacterioph
17	42	34.7	267	1 LIA1 HAEIN	P71392 haemophilus
18	42	34.7	268	1 LIA2 MYCHO	P43056 mycoplasma
19	42	34.7	339	1 LIA2 HAEIN	P4181 haemophilus
20	42	34.7	465	1 TRP2 STRPN	Q54513 streptococc
21	42	34.7	703	1 ARYE MANSE	P44287 manduca sex
22	42	34.7	758	1 PSHA SALTY	P37600 salmonella
23	42	34.7	1017	1 LRGI YEAST	P35688 saccharomyc
24	42	34.7	1193	1 DPOL ADRO4	P87503 human adeno
25	41	33.9	160	1 SPT ENTCL	P39433 enterobacte
26	41	33.9	265	1 PRP2 NPVOP	O10273 ergyia pseu
27	41	33.9	313	1 VGL2 YEAST	P53133 saccharomyc
28	41	33.9	319	1 OBB2 HUMAN	Q36r30 homo sapien
29	41	33.9	319	1 HUTG BACSU	P42088 bacillus su
30	41	33.9	475	1 L14B PRUAR	O24467 prunus arme
31	41	33.9	527	1 YEIP ECOLI	P75785 escherichia
32	41	33.9	561	1 SCW1 SCHPO	O74452 schizosacch
33	41	33.9	645	1 SLT ECOLI	P03810 escherichia

34	41	33.9	645	1 SLT SALTY	P39434 salmonella
35	41	33.9	780	1 YH01 SCHPO	Q10331 schizosacch
36	41	33.9	1017	1 DPOL ADEB2	O72539 bovine aden
37	41	33.9	1188	1 DPOL ADE40	P48311 human adeno
38	41	33.9	5217	1 HPS1 COCCA	Q01886 cochliobolu
39	40.5	33.5	260	1 YFAX ECOLI	P77732 escherichia
40	40.5	33.5	666	1 STI2 KLULA	Q08400 kluyveromyc
41	40	33.1	122	1 Y14K PMV	P20959 papaya mosa
42	40	33.1	141	1 HBA SPFPU	P10059 sphenodon p
43	40	33.1	230	1 FLAH METVO	O06641 methanococc
44	40	33.1	423	1 YE18 SCHPO	O13873 schizosacch
45	40	33.1	548	1 SEAS TOBAC	Q40577 nicotiana t

ALIGNMENTS

RESULT 1
CC27_HUMAN
ID CC27_HUMAN STANDARD; PRT; 824 AA.
AC F30260; C16349; Q96F35;
DT 01-APR-1993 (Rel. 25, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein CDC27hs (Cell division cycle protein 27 homolog) (H-NUC).
GN CDC27.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94052097; PubMed=8234252;
RA Tugendreich S., Boguski M.S., Seldin M., Hieter P.A.;
RT "Linking yeast genetics to mammalian genomes: identification and
RT mapping of the human homolog of CDC27 via the expressed sequence tag
RT (EST) data base."
RL Proc. Natl. Acad. Sci. U.S.A. 90:10031-10035(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95275739; PubMed=7756179;
RA Chen P.L., Ueng Y.C., Durfee T., Chen K.C., Yang-Feng T., Lee W.H.;
RT "Identification of a human homologue of yeast nuc2 which interacts
RT with the retinoblastoma protein in a specific manner."
RL Cell Growth Differ. 6:199-210(1995).
RN [3]
RP SEQUENCE FROM N.A.
RT TISSUE=uterus;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.W., Grumman J.J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBUNIT: INTERACTS WITH RB.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE CDC27/NUC2 FAMILY.
CC -!- SIMILARITY: Contains 9 TPR repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U00001; AAA60471.1; -
 DR EMBL; S78234; AAB34378.1; -
 DR EMBL; BC011656; AAH11656.1; -
 DR PIR; I52835; I52835.
 DR Genew; HGNC:1728; CDC27.
 DR MIM; 116946; -
 DR GO; GO:0005813; C:centrosome; TAS.
 DR GO; GO:0005819; C:spindle; TAS.
 DR GO; GO:0008283; P:cell proliferation; TAS.
 DR GO; GO:0007091; P:mitotic metaphase/anaphase transition; TAS.
 DR InterPro; IPR001440; TPR.
 DR SMART; SM00028; TPR; 8.
 KW Repeat; TPR repeat; Nuclear protein; Polymorphism.
 FT REPEAT 84 114 TPR 1.
 FT REPEAT 115 148 TPR 2.
 FT REPEAT 499 532 TPR 3.
 FT REPEAT 567 600 TPR 4.
 FT REPEAT 602 634 TPR 5.
 FT REPEAT 635 668 TPR 6.
 FT REPEAT 670 702 TPR 7.
 FT REPEAT 704 736 TPR 8.
 FT REPEAT 737 770 TPR 9.
 FT VARIANT 496 496 Y -> H (IN DBSNP:136666).
 FT CONFLICT 319 319 K -> KTRVLQ (IN REF. 3).
 FT CONFLICT 403 403 K -> E (IN REF. 3).
 FT CONFLICT 460 460 MISSING (IN REF. 1).
 FT CONFLICT 715 715 A -> R (IN REF. 1).
 SQ SEQUENCE 824 AA; 91867 MW; E6CBF59C1EF1DCBA CRC64;

Query Match 48.8%; Score 59; DB 1; Length 824;
 Best Local Similarity 54.5%; Pred. No. 0.071; 7; Indels 0; Gaps 0;
 Matches 12; Conservative 3; Mismatches 7;

OY 3 LQLLARCYLSNSQAYGYILK 24
 DB 41 LFLATCYVRSKAYKAYRLK 62

RESULT 2
 BIMA_EMBI STANDARD; PRT; 806 AA.
 AC P17885;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein bima.
 GN BIMA.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R153;
 RC MEDLINE=92121243; PubMed=1770001;
 RA O'Donnell K.L., Osmani A.H., Osmani S.A., Morris N.R.;
 RA "bima" encodes a member of the tetrapeptide repeat family of
 RA proteins and is required for the completion of mitosis in Aspergillus
 RA nidulans.";
 RL J. Cell Sci. 99:711-719(1991).
 CC -!- FUNCTION: REQUIRED FOR THE COMPLETION OF MITOSIS IN ASPERGILLUS

NIDULANS.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- SIMILARITY: BELONGS TO THE CDC27/NUC2 FAMILY.
 CC -!- SIMILARITY: Contains 8 TPR repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X59269; CAA41959.1; -
 DR PIR; A53256; AS3256.
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 7.
 DR SMART; SM00028; TPR; 7.
 KW Cell division; Cell cycle; Mitosis; Repeat; TPR repeat;
 KW Nuclear protein.
 FT REPEAT 76 109 TPR 1.
 FT REPEAT 127 160 TPR 2.
 FT DOMAIN 260 399 BIMA DOMAIN.
 FT REPEAT 513 546 TPR 3.
 FT REPEAT 581 614 TPR 4.
 FT REPEAT 616 648 TPR 5.
 FT REPEAT 649 682 TPR 6.
 FT REPEAT 684 716 TPR 7.
 FT REPEAT 751 784 TPR 8.
 SQ SEQUENCE 806 AA; 89714 MW; F137BDE3A74C0457 CRC64;

Query Match 38.0%; Score 46; DB 1; Length 806;
 Best Local Similarity 56.2%; Pred. No. 10;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 5 LLARCYLSNSQAYSAY 20
 DB 48 LLALCYLQNGQVRAAW 63

RESULT 3
 YG22_HAEIN STANDARD; PRT; 161 AA.
 ID_YG22_HAEIN
 AC P44275; HAEIN
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein H1622 precursor.
 GN H1622.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RC MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clifton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RA Rd.";
 RL Science 269:496-512(1995).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U32835; AAC23271.1; --
DR TIGR; H11622; --
DR Hypothetical protein; Transmembrane, Signal; Complete proteome.
KW SIGNAL 23 OR 21 (POTENTIAL).
FT CHAIN 24 161 HYPOTHETICAL PROTEIN H11622.
FT TRANSMEM 129 149 POTENTIAL.
SQ SEQUENCE .161 AA; 17520 MW; C9303AA13FF25D36 CRC64;

Query Match 37.2%; Score 45; DB 1; Length 161;
Best Local Similarity 50.0%; Pred. No. 2.4; Mismatches 0; Gaps 0;
Matches 8; Conservative 3; Indels 5;

Qy 5 LLARCYLSNSQAYSAY 20
| | | | | : : :
Db 10 LFAACYLPNAVAYALY 25

RESULT 4
YFIM CAEBL STANDARD; PRT; 1459 AA.

AC Q21874;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 166.0 kDa protein R09E10.5 in chromosome IV.
GN R09E10.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Matthews L.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO C.ELEGANS F54D1.6.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; Z70287; CAA94300.1; --
DR PIR; T24088; T24088.
DR WormPep; R09E10.5; CE06287.
DR InterPro; IPR005533; AMOP.
DR InterPro; IPR003886; Nidogen_ext.
DR Pfam; PF03782; AMOP; 1.
DR SMART; SM00723; AMOP; 1.
DR SMART; SM00539; NID0; 1.
KW Hypothetical protein; Transmembrane.
SQ SEQUENCE 1459 AA; 165994 MW; 769014C5599513C4 CRC64;

Query Match 37.2%; Score 45; DB 1; Length 1459;
Best Local Similarity 58.3%; Pred. No. 29;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 9 CYLSNSQAYSAY 20
| | | | | : : :
Db 739 CYMGSSNIYSSY 750

RESULT 5
SAS3_YEAST

ID SAS3_YEAST STANDARD; PRT; 831 AA.
AC P34218;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE SAS3 protein.
GN SAS3 OR YBL052C OR YBL0515 OR YBL0507.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RC MEDLINE=94205266; PubMed=8154187;
RA Scherens B., el Bakkoury M., Vierendeels F., Dubois E., Messenguy F.;
RT "Sequencing and functional analysis of a 32,560 bp segment on the
RT left arm of yeast chromosome II. Identification of 26 open reading
RT frames, including the KIP1 and SEC17 genes."
RL Yeast 9:1355-1371 (1993).
RN [2]
RN CHARACTERIZATION.
RX MEDLINE=96376969; PubMed=8782818;
RA Reifsnnyder C., Lowell J., Clarke A., Pillus L.;
RT "Yeast SAS silencing genes and human genes associated with AML and
RT HIV-1 tat interactions are homologous with acetyltransferases."
RL Nat. Genet. 14:42-49 (1996).
RN [3]
RN ERRATUM.
RP MEDLINE=97285133; PubMed=9140406;
RA Reifsnnyder C., Lowell J., Clarke A., Pillus L.;
RL Nat. Genet. 16:109-109 (1997).
CC -1- FUNCTION: INVOLVED IN SILENCING THE HMR LOCUS.
CC -1- SIMILARITY: BELONGS TO THE MYST (SAS/MOZ) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; Z23261; CAA80794.1; --
DR EMBL; Z35814; CAA84873.1; --
DR PIR; S39835; S39835.
DR SGD; S0000148; SAS3.
DR GO; GO:0016407; P:acetyltransferase activity; IDA.
DR GO; GO:0016568; P:chromatin modification; IMP.
DR GO; GO:0006347; P:chromatin silencing at HML and HMR (sensu S. . .; IMP.
DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.
DR InterPro; IPR002717; MOZ SAS.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF01853; MOZ_SAS; 1.
DR Pfam; PF00096; Zf-C2H2; 1.
KW Zinc-finger.
FT ZN FING 301 323 C2HC-TYPE.
FT DOMAIN 736 831 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 831 AA; 97582 MW; ACF5B1B225CB4A71 CRC64;

Query Match 36.8%; Score 44.5; DB 1; Length 831;
Best Local Similarity 37.0%; Pred. No. 18;
Matches 10; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

Qy 2 NLQLLARCYLSNSQAY----SAYVIL 23
| | | | | : : :
Db 355 NLCLLAKCFINSKTLTYDVERPFIYIL 381

RESULT 6
PSD3 CAEBL STANDARD; PRT; 504 AA.
ID _PSD3 CAEBL
AC Q04908;


```

SQ SEQUENCE 1102 AA; 121516 MW; C5258D714C42PE7A CRC64;
Query Match 36.08; Score 43.5; DB 1; Length 1102;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 13; Conservative 1; Mismatches 5; Indels 7; Gaps 2;

QY 6 LARCYLSN---SOAYSAY---YTLK 24
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
408 LGHCYLQDLDLSRAYSAYRQALYHLK 433

RESULT 8
Y302 STAAW STANDARD; PRT; 266 AA.
AC Q8NYB7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MW0302.
GN MW0302.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN SEQUENCE FROM N.A.
RP MEDLINE=22040717; PubMed=12044378;
RX Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA.";
RL Lancet 359:1819-1827(2002).
CC -!- SIMILARITY: BELONGS TO THE UPF0189 FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; AP004823; BAB94167.1; -
CC EMBL; AP003130; BAB41538.1; -
CC FIRM; G89797; G89797.
CC HAMAP; MF_01205; fused; 1.
CC InterPro; IPR002589; Alpp.
CC Pfam; PF01661; Alpp; 1.
CC SMART; SM00506; Alpp; 1.
CC Hypothetical protein; Complete proteome.
CC DOMAIN 1 79 UNKNOWN.
CC SEQUENCE 266 AA; 30128 MW; 11DBAD2AABA3C083 CRC64;

Query Match 35.5%; Score 43; DB 1; Length 266;
Best Local Similarity 66.7%; Pred. No. 9.1;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VNQLQARCYLS 12
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
182 MNQDLARCYLS 193

RESULT 9
Y325 STAAW STANDARD; PRT; 266 AA.
AC Q99WQ1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein SAV0325/SA0314.
GN SAV0325 OR SA0314.
OS Staphylococcus aureus (strain M50 / ATCC 700699), and
Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878; 158879;

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=M50 / ATCC 700699, and N315;
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus.";
RL Lancet 357:1225-1240(2001).
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE UPF0189
FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; AP003359; BAB56487.1; -
CC EMBL; AP003130; BAB41538.1; -
CC FIRM; G89797; G89797.
CC HAMAP; MF_01205; fused; 1.
CC InterPro; IPR002589; Alpp.
CC Pfam; PF01661; Alpp; 1.
CC SMART; SM00506; Alpp; 1.
CC Hypothetical protein; Complete proteome.
CC DOMAIN 1 79 UNKNOWN.
CC SEQUENCE 266 AA; 30109 MW; 5B57528F6BC4F16B CRC64;

Query Match 35.5%; Score 43; DB 1; Length 266;
Best Local Similarity 66.7%; Pred. No. 9.1;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VNQLQARCYLS 12
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
182 MNQDLARCYLS 193

RESULT 10
WTBI SCOLI
ID WTBI_ECOLI STANDARD; PRT; 325 AA.
AC P00472;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Modification methylase EcoRI (EC 2.1.1.72) (Adenine-specific
methyltransferase EcoRI (M.EcoRI)).
GN ECORIM.
OS Escherichia coli.
OG Plasmid pMB1, and Plasmid pMB4.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN SEQUENCE FROM N.A.
RP PLASMID=pMB1;
RX MEDLINE=81117320; PubMed=6257703;
RA Greene P.J., Gupta M., Boyer H.W., Brown W.E., Rosenberg J.M.;
RT "Sequence analysis of the DNA encoding the Eco RI endonuclease and
methylase.";
RL J. Biol. Chem. 256:2143-2153(1981).
RN SEQUENCE FROM N.A.
RP STRAIN=K12 / C600; PLASMID=pMB4;
RX MEDLINE=81117318; PubMed=6257701;

```

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; X15828; CAC16944.1; --
InterPro; IPR001988; Caulimo_coat.
InterPro; IPR001878; Znf_CCHC.
Pfam; PF00098; zf-CCHC; 1.
PRINTS; PR00939; C2HCZNFINGER.
PRINTS; PR00221; CAULIMOCOAT.
SMART; SM00343; ZnF_C2HC; 1.
PROSITE; PS0158; ZF_CCHC; 1.
Coat protein; zinc-finger.
Zn_Fing 381
SEQUENCE 441 AA; 52102 MW; 07244CD2181CFPAFF CRC64;
CCHC-TYPE.

Query Match 35.5%; Score 43; DB 1; Length 441;
Best Local Similarity 39.1%; Pred. No. 16;
Matches 9; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

QY 1 VNQLRLARCYLSN--SQAYSAYY 21
:|::|||-:-:||:
Db 216 INLEICNMCYLENFLCEFOGRYY 238

RESULT 12
TI60_HUMAN STANDARD; PRT; 513 AA.
ID AC Q92993; O95624; Q13430; Q9BMK7;
DT 15-JUL-1998 (Rel. 36, Created)
DT DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 60 Kda Tat interactive protein (Tip60) (HIV-1 Tat interactive
protein) (cPLA(2) interacting protein).
GN HTATIP OR TIP60.
OS Homo sapiens (Human).
OC Buckyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP RP
RX MEDLINE=96182937; PubMed=8607285;
RA Kamine J., Elangovan B., Subramanian T., Coleman D.,
Chinnadurai G.
RT Identification of a cellular protein that specifically interacts
RW with the essential cysteine region of the HIV-1 Tat transactivator.";
RL Virology 216:357-366(1996).
[2]
RN SEQUENCE FROM N.A. (ISOFORM 2), INTERACTION WITH PLA2G4A, AND
SUBCELLULAR LOCATION.
RP RP
RC TISSUE=Lymphoblast,
TX MEDLINE=96182937; PubMed=8607285;
RX Kamine J., Elangovan B., Subramanian T., Coleman D.,
Chinnadurai G.
RA Sheridan A.M., Force T., Yoon H.J., O'Leary E., Choukroun G.,
Taheri M.R., Bonventre J.V.;
RT "PLIIF, a novel splice variant of Tip60, interacts with group IV
RT cytosolic phospholipase A(2), induces apoptosis, and potentiates
RT prostaglandin production."
RL Mol. Cell. Biol. 21:4470-4481(2001).
[3]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RP RP
RC TISSUE=Cervix;
RX MEDLINE=22386257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
R Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
ALtschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
DIatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J.J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4].
RP INTERACTION WITH EDNRA.
RX PubMed=11262386;
RA Lee H.-J., Chun M., Kandrak K.V.;
RT "Tip60 and HDAC7 interact with the endothelin receptor a and may be
RT involved in downstream signaling.";
RL J. Biol. Chem. 276:16597-16600(2001).
RN [5].
RP INTERACTION WITH HDAC7.
RX PubMed=12551922;
RA Xiao H., Chung J., Kao H.-Y., Yang Y.-C.;
RT "Tip60 is a co-repressor for STAT3";
RL J. Biol. Chem. 278:11197-11204(2003).
CC -!- FUNCTION: Binds to the TAT protein of the human immunodeficiency
CC virus (HIV). Specific binding of Tip60 to TAT might be an
CC important feature for efficient TAT transactivation of HIV gene
CC expression.
CC -!- SUBUNIT: Interacts with HIV1 TAT, PLA2G4A/CPLA2, EDNRA and HDAC7.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q22993-1; Sequences=Displayed;
CC Name=2; Synonyms=PLIP;
CC IsoId=Q22993-2; Sequences=VSP 007438;
CC -!- SIMILARITY: BELONGS TO THE MYST (SAS/MOZ) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U74667; AAB18236.1; -;
DR EMBL; U40989; AAB02683.1; ALT_INIT.
DR EMBL; U67734; AAD00163.1; -;
DR EMBL; BC000166; AAB00166.1; -;
DR Genew; HGNC:5275; HTATIP.
DR MIM; 601409; -;
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003713; F:transcription co-activator activity; TAS.
DR GO; GO:0008366; P:transcription from Pol II promoter; TAS.
DR InterPro; IPR000953; Chromo.
DR InterPro; IPR002717; MOZ SAS.
DR Pfam; PF01853; MOZ SAS; 1.
DR SMART; SM00298; CHROMO; 1.
KW Nuclear protein; Zinc-finger; Alternative splicing.
FT ZN FING 261 283
FT VARSPPLIC 96 147
FT Missing (in isoform 2).
FT /FTID=VSP_007438.
FT CONFLICT 382 382 G -> R (IN REF. 1).
FT SEQUENCE 513 AA; 58581 MW; 63724F5E10B957D5 CRC64;

Query Match 35.5%; Score 43; DB 1; Length 513;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 NLCLARCYLSNSQAY 17
Db 315 NLCLAKCFDHRKTY 330
RESULT 13
VB04 VACCC
ID VB04 VACCC STANDARD; PRT; 558 AA.
AC P21001;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein B4.
GN B4.
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10249;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021027; PubMed=2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paolletti E.;
RT "The complete DNA sequence of vaccinia virus.";
RL Virology 179:247-266(1990).
RN [2]
RP COMPLETE GENOME.
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paolletti E.;
RT "Appendix to 'the complete DNA sequence of vaccinia virus.'";
RL Virology 179:517-563(1990).
CC -!- SIMILARITY: Contains 7 ANK repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M35027; AAA48200.1; -;
DR PIR; C42526; C42526.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 6.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR Late protein; Repeat; ANK repeat.
FT REPEAT 65 95 ANK 1.
FT REPEAT 169 205 ANK 2.
FT REPEAT 209 239 ANK 3.
FT REPEAT 243 272 ANK 4.
FT REPEAT 276 304 ANK 5.
FT REPEAT 339 368 ANK 6.
FT REPEAT 372 401 ANK 7.
SQ SEQUENCE 558 AA; 65272 MW; ED672F454DB7FC76 CRC64;
Query Match 35.5%; Score 43; DB 1; Length 558;
Best Local Similarity 22.7%; Pred. No. 21;
Matches 5; Conservative 7; Mismatches 10; Indels 0; Gaps 0;
QY 3 LQLARCYLSNSQAYSAYILK 24
Db 456 IDIMKRCYIKNTVFQVFCIK 477
RESULT 14
VB04 VACCV
ID VB04 VACCV STANDARD; PRT; 558 AA.
AC P24769;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 2, 2004, 09:20:42 ; Search time 177 Seconds
(without alignments)
34.990 Million cell updates/sec

Title: US-10-036-492-6

Perfect score: 121
Sequence: 1 VNLQRLARCVLSNQAYSAYILK 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_nhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	75.2	471	10 Q8W4M8	Q8W4M8 arabidopsis
2	91	75.2	744	10 Q8LQUG	Q8LQUG arabidopsis
3	59	48.8	394	11 Q8BYJ1	Q8BYJ1 mus musculus
4	59	48.8	399	11 Q8S568	Q8S568 mus musculus
5	53	43.8	796	16 Q8X715	Q8X715 escherichia
6	52	43.0	300	12 Q9YUQ5	Q9YUQ5 turkey aden
7	49	40.5	146	5 Q9BPL6	Q9BPL6 metagonimus
8	49	40.5	262	16 Q8DPU8	Q8DPU8 streptococc
9	49	40.5	875	5 Q24021	Q24021 drosophila
10	49	40.5	900	5 Q9V637	Q9V637 drosophila
11	48	39.7	501	5 Q19498	Q19498 caenorhabdi
12	47	38.8	298	16 Q97GR8	Q97GR8 clostridium
13	47	38.8	758	5 Q8IAY2	Q8IAY2 plasmodium
14	46.5	38.4	7149	5 Q8IM09	Q8IM09 plasmodium
15	46	38.0	173	17 Q58958	Q58958 pyrococcus
16	46	38.0	300	10 Q39979	Q39979 hyoscyamus

17	46	38.0	332	10 Q8L7I9	Q8L7I9 arabidopsis
18	46	38.0	437	10 Q9MID6	Q9MID6 arabidopsis
19	46	38.0	441	4 Q8N7G5	Q8N7G5 homo sapien
20	46	38.0	453	10 Q9FVL2	Q9FVL2 lycopersico
21	46	38.0	454	10 Q8LK91	Q8LK91 arabidopsis
22	46	38.0	458	10 Q8L727	Q8L727 arabidopsis
23	46	38.0	462	10 Q8L7S2	Q8L7S2 arabidopsis
24	46	38.0	520	10 Q39978	Q39978 hyoscyamus
25	46	38.0	549	10 Q9XJ20	Q9XJ20 solanum tub
26	46	38.0	549	10 Q9XJ25	Q9XJ25 solanum tub
27	46	38.0	550	10 Q9SB70	Q9SB70 solanum tub
28	46	38.0	551	10 Q9ZTQ6	Q9ZTQ6 solanum tub
29	46	38.0	552	5 Q8MQB6	Q8MQB6 caenorhabdi
30	46	38.0	553	10 Q9ATN6	Q9ATN6 capsicum an
31	46	38.0	556	10 Q9XJ32	Q9XJ32 solanum tub
32	46	38.0	557	10 Q9ZTQ7	Q9ZTQ7 solanum tub
33	46	38.0	557	10 Q9ZTQ8	Q9ZTQ8 solanum tub
34	46	38.0	574	5 Q95QT8	Q95QT8 caenorhabdi
35	46	38.0	582	6 Q9GM29	Q9GM29 macaca fasc
36	46	38.0	724	6 Q95JZ1	Q95JZ1 macaca fasc
37	46	38.0	882	4 Q8NA02	Q8NA02 homo sapien
38	46	38.0	1145	12 Q9IV56	Q9IV56 olive laten
39	46	38.0	1164	17 Q8PX58	Q8PX58 methanosarc
40	46	38.0	1320	4 Q8NDW8	Q8NDW8 homo sapien
41	46	38.0	1772	5 Q9W040	Q9W040 drosophila
42	45	37.2	455	3 Q88231	Q88231 saccharomyc
43	45	37.2	508	5 Q9BL61	Q9BL61 caenorhabdi
44	45	37.2	527	16 Q8F0N0	Q8F0N0 leptospira
45	45	37.2	759	2 Q93M42	Q93M42 streptococc

ALIGNMENTS

RESULT 1

Q8W4M8 PRELIMINARY; PRT; 471 AA.

AC Q8W4M8;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE CDC27/NUC2-like protein.
GN AT2G20000, T2G17.20 OR AT2G20000.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosida II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]_
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY062470; AAL32548.1; -
DR EMBL; AY128780; AAM91180.1; -
DR InterPro; IPR001440; TPR.
SQ SEQUENCE 471 AA; 51636 MW; 4F0D2BA57F94043 CRC64;

Query Match 75.2%; Score 91; DB 10; Length 471;
Best Local Similarity 75.0%; Pred. NO. 1.8e-06;

Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VNQLLARCYSNSOAYSAYILK 24
 Db 36 VNQLLATSYLQNNQAYSAYHLK 59

RESULT 2

Q8LGU6 PRELIMINARY; PRT; 744 AA.
 AC Q8LGU6;
 DT 01-OCT-2002 (TREMELrel. 22, Created)
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE HOBBIT protein.
 GN HBT.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Green siliques;
 RA Blilou I., Frugier F., Folmer S., Serralbo O., Willemse V.,
 RA Wolkfelt H., Eloy N., Ferreira P., Weisbeek P., Scheres B.J.G.;
 RT "The Arabidopsis HOBBIT gene encodes a CDC27 homologue that links the
 RT plant cell cycle to progression of cell differentiation."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; AJ487669; CAD31951.1; -
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 2.
 DR SMART; SM00028; TPR; 8.
 SQ SEQUENCE 744 AA; 83068 MW; 1CB1719C75469861 CRC64;

Query Match 75.2%; Score 91; DB 10; Length 744;
 Best Local Similarity 75.0%; Pred. No. 2.9e-06;
 Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VNQLLARCYSNSOAYSAYILK 24
 Db 36 VNQLLATSYLQNNQAYSAYHLK 59

RESULT 3

Q8BYJ1 PRELIMINARY; PRT; 394 AA.
 AC Q8BYJ1;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Protein CDC27HS (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK039377; BAC30331.1; -
 FT NON TER 394
 SQ SEQUENCE 394 AA; 42739 MW; 9693509545DDE25D CRC64;

Query Match 48.8%; Score 59; DB 11; Length 394;
 Best Local Similarity 54.5%; Pred. No. 0.26;
 Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LQLLARCYSNSOAYSAYILK 24
 Db 41 LFLLATCYRSGKAYKAYILK 62

RESULT 4

Q8R568 PRELIMINARY; PRT; 399 AA.
 AC Q8R568;
 DT 01-JUN-2002 (TREMELrel. 21, Created)
 DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
 DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
 DE Similar to cell division cycle 27.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; BC023187; AAH23187.1; -
 DR InterPro; IPR001440; TPR.
 KW Cell division.
 SQ SEQUENCE 399 AA; 43328 MW; EA585749473077B2 CRC64;

Query Match 48.8%; Score 59; DB 11; Length 399;
 Best Local Similarity 54.5%; Pred. No. 0.26;
 Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LQLLARCYSNSOAYSAYILK 24
 Db 41 LFLLATCYRSGKAYKAYILK 62

RESULT 5

Q8X7I5 PRELIMINARY; PRT; 796 AA.
 AC Q8X7I5;
 DT 01-MAR-2002 (TREMELrel. 20, Created)
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Alpha replication protein of prophage CP-933I (Putative DNA
 DE primase).
 GN Z0339 OR ECS0303.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
 RL Nature 409:529-533 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurckawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Tasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12."
 RL DNA Res. 8:11-22 (2001).
 DR EMBL; AE005204; AAC54598.1; -

```

DR EMBL; AF002551; BAB33726.1; ALT_INIT.
DR InterPro; IPR006500; Primase_C.
DR InterPro; IPR006154; Toprim_sub.
DR SMART; SM00493; TOPRIM; 1.
DR TIGRFAMs; TIGR01613; primase_Cterm; 1.
KW Complete proteome.
SQ SEQUENCE 796 AA; 87237 MW; 5ABEBE08B6020A672 CRC64;

Query Match 43.8%; Score 53; DB 16; Length 796;
Best Local Similarity 35.7%; Pred. No. 5.3;
Matches 10; Conservative 8; Mismatches 6; Indels 4; Gaps 1;

QY 1 VNQLQL---ARCYLSNSQAYSAYILK 24
DB 193 VNQLQLINANGKCFILKGGQVNAFLVE 220

RESULT 6
Q9YUQ5 PRELIMINARY; PRT; 300 AA.
AC Q9YUQ5
AC Q9YUQ5
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE E3 protein.
DE E3
OS Turkey adenovirus 3.
OS Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Siadenovirus.
OS NCBI_TaxID=41678;
OX [1]
RN
RP SEQUENCE FROM N.A.
RP MEDLINE=99009384; PubMed=9791022;
RX Pitcovski J., Muallem M., Rei-Koren Z., Krispel S., Gallili G.,
RA Michael A., Goldberg D.;
RT "The complete DNA sequence and genome organization of the avian
RT adenovirus, hemorrhagic enteritis virus.";
RL Virology 249:307-315(1998).
RN [2]
RN
RP SEQUENCE FROM N.A.
RP Pitcovski J., Muallem M., Rei-Koren Z., Krispel S., Gallili G.,
RA Michael A., Goldberg D.;
RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RL
RL EMBL; AF074946; AAC64537.1; --
SQ SEQUENCE 300 AA; 34202 MW; 9922C0B4A8B62E52 CRC64;

Query Match 43.0%; Score 52; DB 12; Length 300;
Best Local Similarity 41.7%; Pred. No. 2.7;
Matches 10; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 VNQLQLARCYLSNSQAYSAYILK 24
DB 216 VNFPFLCSNCFKGNDCSCAFYICK 239

RESULT 7
Q9BPL6 PRELIMINARY; PRT; 146 AA.
AC Q9BPL6
AC Q9BPL6
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cysteine proteinase (Fragment).
DE CP4.
OS Metagonimus yokogawai.
OS Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
OC Opisthorchiida; Opisthorchiata; Opisthorchioidea; Heterophyidae;
OC Metagonimus.
OX NCBI_TaxID=84529;
RN [1]
RN
RP SEQUENCE FROM N.A.
RP Park H., Yun H.C., Kim K.Y., Park S.Y., Park S.K.;
RT "Metagonimus yokogawai cysteine proteinase (MYCP4) mRNA.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

```

RA Tugendreich S., Johns Hopkins University.
 RL Thesis (1995); AA57340.1; -.
 DR EMBL; U18298; AA57340.1; -.
 DR FlyBase; FBgn0012058; Cdc27.
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 6.
 DR SMART; SM00028; TPR; 5.
 SQ SEQUENCE 875 AA; 97658 MW; 37662A8843FF39FD CRC64;

Query Match 40.5%; Score 49; DB 5; Length 875;
 Best Local Similarity 45.0%; Pred. No. 27;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 5 LLARCYLSNSQAYSAYVILK 24
 ||||| : : : : :
 Db 42 LLATSYFRSNQVHQAYWLK 61

RESULT 10

Q9VS37 Q9VS37 PRELIMINARY; PRT; 900 AA.
 AC Q9VS37;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE CDC27 protein.
 GN Drosophila melanogaster (Fruit fly).
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Anandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazei R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Malshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasgaman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";

RL Science 287:2185-2195(2000).
 DR EMBL; AB003559; AAF50592.1; -.
 DR FlyBase; FBgn0012058; Cdc27.
 DR InterPro; IPR001440; TPR.
 DR Pfam; PF00515; TPR; 8.
 DR SMART; SM00028; TPR; 7.
 SQ SEQUENCE 900 AA; 101284 MW; 363E33789A783464 CRC64;

Query Match 40.5%; Score 49; DB 5; Length 900;
 Best Local Similarity 45.0%; Pred. No. 27;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 5 LLARCYLSNSQAYSAYVILK 24
 ||||| : : : : :
 Db 42 LLATSYFRSNQVHQAYWLK 61

RESULT 11

Q19498 Q19498 PRELIMINARY; PRT; 501 AA.
 AC Q19498;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 56.8 kDa protein.
 GN F16H11.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peleoderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=93069613; PubMed=9851916;
 RA None;
 RT Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.;
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wu X.;
 RT "The sequence of C. elegans cosmid F16H11.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U55376; AA98007.1; -.
 DR WormPep; F16H11.1; CE04396.
 KW Hypothetical protein.
 SQ SEQUENCE 501 AA; 56826 MW; E21783C13CE3B95E CRC64;

Query Match 39.7%; Score 48; DB 5; Length 501;
 Best Local Similarity 45.0%; Pred. No. 21;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 3 LQILARCYLSNSQAYSAYYI 22
 ||||| : : : : :
 Db 287 LYMLSLRYINISQVFFFYI 306

RESULT 12

Q97GR8 Q97GR8 PRELIMINARY; PRT; 298 AA.
 AC Q97GR8;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Predicted nucleocytoplasmic transferase.
 GN CAC2298.

OS Clostridium acetobutylicum
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson K., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatsov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 bacterium Clostridium acetobutylicum";
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL; AF007730; AAK80254.1; -.
 DR InterPro; IPR002934; NTP.transf.
 DR Pfam; PF01909; NTP.transf.; 1.
 KW Transferase; Complete proteome.
 SQ SEQUENCE 298 AA; 35000 MW; 9986BOAB8FEED303 CRC64;
 Query Match 38.8%; Score 47; DB 16; Length 298;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 6 LARCYLSNQAYSAYY 21
 Db 142 LCKKYLNNKRIYTAY 157
 RESULT 13 -
 Q8IAY2 PRELIMINARY; PRT; 758 AA.
 ID Q8IAY2
 AC Q8IAY2
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN MAL8PI.88.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Seeger K., Murphy L., Harris D., Berriman M., Pain A., Hall N.,
 RA Quail M., Barrell B.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL844507; CAD51228.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 758 AA; 90296 MW; B9C395E9D9103492 CRC64;
 Query Match 38.8%; Score 47; DB 5; Length 758;
 Best Local Similarity 59.2%; Pred. No. 49;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 8 RCYLSNSQAYSAY 20
 Db 121 RCYLSNSLYSKY 133
 RESULT 14
 Q8IM09 PRELIMINARY; PRT; 7149 AA.
 ID Q8IM09
 AC Q8IM09;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN PF14_0084.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN (1)

RP SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RX MEDLINE=22555705; PubMed=12368864;
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eissen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA Chan M.-S., Nene J., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
 RA Partea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
 RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.W., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffmann S.D., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrell B.;
 RT "Genome sequence of the human malaria parasite Plasmodium
 falciparum";
 RL Nature 419:498-511(2002).
 DR EMBL; AB014817; AAN36596.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 7149 AA; 847876 MW; 4E3DD7D9FE1199B5 CRC64;
 Query Match 38.4%; Score 46.5; DB 5; Length 7149;
 Best Local Similarity 38.1%; Pred. No. 6.4e+02;
 Matches 8; Conservative 6; Mismatches 4; Indels 3; Gaps 1;
 QY 2 NLCLLARCYSLSQAYSAYYI 22
 Db 2496 NVRL--CYNNSEGEYKHYV 2513
 RESULT 15
 O58958 PRELIMINARY; PRT; 173 AA.
 ID O58958
 AC O58958;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein PH1225.
 GN PH1225.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3";
 RL DNA Res. 5:555-76(1998).
 DR EMBL; AP000005; BAA30325.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 173 AA; 19986 MW; 7D7FD7D9BB798B1B CRC64;
 Query Match 38.0%; Score 46; DB 17; Length 173;
 Best Local Similarity 45.8%; Pred. No. 15;
 Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
 QY 1 VNLQLIARCYSLSQAYSAYYILK 24
 Db 62 VGVILWAICYFQVSKAYKAGLKX 85
 Search completed: February 2, 2004, 12:50:22
 Job time : 180 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 2, 2004, 15:40:06 ; Search time 240 Seconds
(without alignments)
364.456 Million cell updates/sec

Title: US-10-036-492-6

Perfect score: 121

Sequence: 1 VNLQLARCYLSNQAYVILK 24

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10036492/runat_02022004_085035_13526/app_query.fasta_1.199
-DB=Published Applications NA -QMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USREP=US10036492@cgn 1.1 221 @runat_02022004_085035_13526
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description

1	121	100.0	2401	14	US-10-036-492-14	Sequence 14, Appl
2	121	100.0	2434	14	US-10-036-492-9	Sequence 27, Appl
3	121	100.0	2512	14	US-10-036-492-27	Sequence 15, Appl
4	91	75.2	2220	14	US-10-036-492-15	Sequence 17374, A
5	61	50.4	419	11	US-09-918-995-17374	Sequence 101808,
6	53	43.8	2318	13	US-10-027-632-101808	Sequence 101809,
7	53	43.8	2318	13	US-10-027-632-101809	Sequence 101808,
8	53	43.8	2318	14	US-10-027-632-101808	Sequence 101809,
9	53	43.8	12886	15	US-10-027-632-101809	Sequence 14, Appl
10	53	43.8	12886	15	US-10-114-170-14	Sequence 192115,
11	52	43.0	585	13	US-10-027-632-192114	Sequence 192115,
12	52	43.0	585	13	US-10-027-632-192115	Sequence 192115,
13	52	43.0	585	14	US-10-027-632-192115	Sequence 192115,
14	52	43.0	585	14	US-10-027-632-192115	Sequence 26594, A
15	52	43.0	722	13	US-10-027-632-26594	Sequence 26594, A
16	52	43.0	722	14	US-10-027-632-26594	Sequence 258286,
17	52	43.0	1094	13	US-10-027-632-258286	Sequence 258287,
18	52	43.0	1094	13	US-10-027-632-258287	Sequence 258288,
19	52	43.0	1094	13	US-10-027-632-258288	Sequence 258286,
20	52	43.0	1094	14	US-10-027-632-258286	Sequence 258287,
21	52	43.0	1094	14	US-10-027-632-258287	Sequence 258288,
22	52	43.0	1094	14	US-10-027-632-258288	Sequence 522, App
23	52	43.0	1987	8	US-08-781-986A-522	Sequence 307, App
24	52	43.0	5771	10	US-09-833-381-307	Sequence 358, App
25	52	43.0	325348	12	US-10-085-117-358	Sequence 2184, App
26	51	42.1	383	10	US-09-878-574-2184	Sequence 555, App
27	51	42.1	412	10	US-09-867-701-555	Sequence 73366, A
28	51	42.1	441	13	US-10-027-632-73366	Sequence 73366, A
29	51	42.1	441	14	US-10-027-632-73366	Sequence 103, App
30	51	42.1	502	10	US-09-919-580-103	Sequence 50, Appl
31	51	42.1	505	10	US-09-998-598-50	Sequence 313046,
32	51	42.1	562	13	US-10-027-632-313046	Sequence 313047,
33	51	42.1	562	13	US-10-027-632-313047	Sequence 313046,
34	51	42.1	562	14	US-10-027-632-313046	Sequence 313047,
35	51	42.1	562	14	US-10-027-632-313047	Sequence 27710, A
36	51	42.1	698	13	US-10-027-632-27710	Sequence 27710, A
37	51	42.1	698	13	US-10-027-632-27710	Sequence 27710, A
38	51	42.1	698	14	US-10-027-632-27710	Sequence 359, App
39	51	42.1	1754	10	US-09-998-598-359	Sequence 32, Appl
40	51	42.1	118067	16	US-10-081-327-32	Sequence 1102, App
41	51	42.1	2007	12	US-10-108-260A-1102	Sequence 38093, A
42	50.5	41.7	487	12	US-10-242-535A-38093	Sequence 93614, A
43	50	41.3	487	12	US-10-242-535A-38093	Sequence 93615, A
44	50	41.3	554	13	US-10-027-632-93614	
45	50	41.3	554	13	US-10-027-632-93615	

ALIGNMENTS

RESULT 1
US-10-036-492-14
; Sequence 14, Application US/10036492
; Publication No. US20020164757A1
; GENERAL INFORMATION:
; APPLICANT: HEMERLY, ADRIANA
; APPLICANT: FERREIRA, PAULO
; APPLICANT: ROMEAUS, STEPHANE
; TITLE OF INVENTION: PLANT DNA REPLICATION MODULATING PROTEINS
; FILE REFERENCE: 217943USOXCONT
; CURRENT APPLICATION NUMBER: US/10/036,492
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: EP 99202214.5
; PRIOR FILING DATE: 1999-07-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 2401
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-036-492-14

Alignment Scores: 6.26e-10 Length: 2401
Pred. No.: 2401

```
Score: 121.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-036-492-6 (1-24) x US-10-036-492-14 (1-2401)
QY 1 ValAsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 109 GTGAACCTGCAATTTGATCCAGGTGTTACTTGAGTAAACAGTCAAGCTTATAGTGCATAT 168
QY 21 TyrIleLeuLys 24
Db 169 TATATCCTTAAA 180

RESULT 2
US-10-036-492-9
; Sequence 9, Application US/10036492
; Publication No. US20020164757A1
; GENERAL INFORMATION:
; APPLICANT: HEMERLY, ADRIANA
; APPLICANT: FERREIRA, PAULO
; APPLICANT: ROMBAUTS, STEPHANE
; TITLE OF INVENTION: PLANT DNA REPLICATION MODULATING PROTEINS
; FILE REFERENCE: 217943USOXCONT
; CURRENT APPLICATION NUMBER: US/10/036,492
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: EP 99202214.5
; PRIOR FILING DATE: 1999-07-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-036-492-9

Alignment Scores:
Pred. No.: 6.37e-10 Length: 2434
Score: 121.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-036-492-6 (1-24) x US-10-036-492-9 (1-2434)
QY 1 ValAsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 109 GTGAACCTGCAATTTGATCCAGGTGTTACTTGAGTAAACAGTCAAGCTTATAGTGCATAT 168
QY 21 TyrIleLeuLys 24
Db 169 TATATCCTTAAA 180

RESULT 3
US-10-036-492-27/c
; Sequence 27, Application US/10036492
; Publication No. US20020164757A1
; GENERAL INFORMATION:
; APPLICANT: HEMERLY, ADRIANA
; APPLICANT: FERREIRA, PAULO
; APPLICANT: ROMBAUTS, STEPHANE
; TITLE OF INVENTION: PLANT DNA REPLICATION MODULATING PROTEINS
; FILE REFERENCE: 217943USOXCONT
; CURRENT APPLICATION NUMBER: US/10/036,492
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: EP 99202214.5
; PRIOR FILING DATE: 1999-07-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 27

US-10-036-492-6 (1-24) x US-10-036-492-15 (1-2220)
QY 1 ValAsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 106 GTTAATTTGCAGCTATTAGCCAGCAGCTACCTGCAGATAATCAAGCTTACAGTGCATAT 165
QY 21 TyrIleLeuLys 24
Db 166 CATCTGCTAAG 177

RESULT 5
US-09-918-995-17374
; Sequence 17374, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
```

```

; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17374
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-17374
Alignment Scores:
Pred. No.: 0.616 Length: 419
Score: 61.00 Matches: 12
Percent Similarity: 68.18% Conservative: 3
Best Local Similarity: 54.55% Mismatches: 7
Query Match: 50.41% Indels: 0
DB: 11 Gaps: 0

US-10-036-492-6 (1-24) x US-09-918-995-17374 (1-419)
Qy 3 LeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyrIle 22
Db 201 TTGCTTTTACTGCAACCTGTTATTACCGCTCAGGAGGCATATAAGCATATAGACTC 260
Qy 23 LeuIys 24
Db 261 TTGAAA 266

RESULT 6
US-10-027-632-101808
; Sequence 101808, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101809
; LENGTH: 2318
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-101808
Alignment Scores:
Pred. No.: 124 Length: 2318
Score: 53.00 Matches: 9
Percent Similarity: 68.18% Conservative: 6
Best Local Similarity: 40.91% Mismatches: 7
Query Match: 43.80% Indels: 0
DB: 13 Gaps: 0

US-10-036-492-6 (1-24) x US-10-027-632-101808 (1-2318)
Qy 2 AsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyr 21
Db 690 AATATGCAATGTTGAGTCTGTGCTCTTAACTCACTCATCTGATGAAAAATCATAT 749
Qy 22 IleLeu 23
Db 750 TTACTT 755

RESULT 8
US-10-027-632-101808
; Sequence 101808, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101808
; LENGTH: 2318
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-101808
Alignment Scores:
Pred. No.: 124 Length: 2318
Score: 53.00 Matches: 9
Percent Similarity: 68.18% Conservative: 6
Best Local Similarity: 40.91% Mismatches: 7
Query Match: 43.80% Indels: 0
DB: 13 Gaps: 0

US-10-036-492-6 (1-24) x US-10-027-632-101808 (1-2318)
Qy 2 AsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyr 21
```


PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 101808
LENGTH: 2318
TYPE: DNA
ORGANISM: Human
US-10-027-632-101808

Alignment Scores:
Pred. No.: 124
Score: 53.00
Percent Similarity: 68.18%
Best Local Similarity: 40.91%
Query Match: 43.80%
DB: 14
Length: 2318
Matches: 9
Conservative: 6
Mismatches: 7
Indels: 0
Gaps: 0

US-10-036-492-6 (1-24) x US-10-027-632-101808 (1-2318)

Qy 2 AsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyr 21
Db 690 AATATGCAATGTTGAGTCGTGCTTACAACTCATCTCGCATGAAATCATTTAT 749

Qy 22 IleLeu 23
Db 750 TTACTT 755

RESULT 9
US-10-027-632-101809
Sequence 101809, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 101809
LENGTH: 2318
TYPE: DNA
ORGANISM: Human
US-10-027-632-101809

Alignment Scores:
Pred. No.: 124
Score: 53.00
Percent Similarity: 68.18%
Best Local Similarity: 40.91%
Query Match: 43.80%
Length: 2318
Matches: 9
Conservative: 6
Mismatches: 7
Indels: 0

DB: 14 Gaps: 0
US-10-036-492-6 (1-24) x US-10-027-632-101809 (1-2318)

Qy 2 AsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyr 21
Db 690 AATATGCAATGTTGAGTCGTGCTTACAACTCATCTCGCATGAAATCATTTAT 749

Qy 22 IleLeu 23
Db 750 TTACTT 755

RESULT 10
US-10-114-170-14
Sequence 14, Application US/10114170
Publication No. US20030023075A1
GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. US20030023075A1 Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/114.170
FILING DATE: 01-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/453,702
FILING DATE: 03-DEC-1999
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 12886
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-10-114-170-14

Alignment Scores:
Pred. No.: 1.17e+03
Score: 53.00
Percent Similarity: 64.29%
Best Local Similarity: 35.71%
Query Match: 43.80%
DB: 15
Length: 12886
Matches: 10
Conservative: 8
Mismatches: 6
Indels: 4
Gaps: 1

US-10-036-492-6 (1-24) x US-10-114-170-14 (1-12886)

Qy 1 ValAsnLeuGlnLeuLeu-----AlaArgCysTyrLeuSerAsnSerGlnAla 16

```

Db      9431 GTAATCTGCAGTTTAATCAACGCTAAATGGGGGAARATGCTTCCTTAAAGCGGTACAGTT 9490
Qy      17 TyrSerAlaTyrrYlleLeuLys 24
Db      9491 AAGAATGCCTTTACCTGGTTGAA 9514

RESULT 11
US-10-027-632-199214
; Sequence 199214, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19214
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-199214

Alignment Scores:
Pred No.:          30              Length:       585
Score:            52.00           Matches:      11
Percent Similarity: 68.42%        Conservative:   2
Best Local Similarity: 57.89%     Mismatches:    2
Query Match:      42.98%         Indels:        4
DB:               13             Gaps:          1

US-10-036-492-6 (1-24) x US-10-027-632-199214 (1-585)

Qy      8 ArgCysrYrLeuSerAsn-----SerGlnAlaTyrrSerAlaTyrrYlle 22
Db      166 AAATGTTATTATCAATGCTTACAACATATCAAACACTATTTCGTGCAATATTATA 222

RESULT 12
US-10-027-632-199215
; Sequence 199215, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24

```

```
RESULT 14
US-10-027-632-199215
; Sequence 199215, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 199215
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-199215

Alignment Scores:
Pred. No.: 30 Length: 585
Score: 52.00 Matches: 11
Percent Similarity: 88.42% Conservative: 2
Best Local Similarity: 57.89% Mismatches: 2
Query Match: 42.98% Indels: 4
DB: 14 Gaps: 1

US-10-036-492-6 (1-24) x US-10-027-632-199215 (1-585)
Qy 8 ArgCysTyrLeuSerAsn-----SerGlnAlaTyrSerAlaTyrTyrIle 22
Db 166 AAATGCTATTATCAATGCTTACACACATATCAACCTATTTCGTCATATTATATA 222

RESULT 15
US-10-027-632-26594
; Sequence 26594, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26594
; LENGTH: 722
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-26594

Alignment Scores:
Pred. No.: 39.5 Length: 722
Score: 52.00 Matches: 13
Percent Similarity: 82.61% Conservative: 6
Best Local Similarity: 56.52% Mismatches: 1
Query Match: 42.98% Indels: 3
DB: 13 Gaps: 1

US-10-036-492-6 (1-24) x US-10-027-632-26594 (1-722)
Qy 2 AsnLeuGlnLeuAlaArgCysTyr-LeuSerAsnSerGlnAlaTyrSerAlaTyr 21
Db 611 TCACACAGTTACTC-----TGTATGCTATCAACACACTAGATCTTATTCATT 664

Qy 21 rIleLeu 23
Db 665 CATTATA 671

Search completed: February 2, 2004, 19:41:51
Job time : 246 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 2, 2004, 13:26:34 ; Search time 163 Seconds
(without alignments)

64.989 Million cell updates/sec

Title: US-10-036-492-6

Perfect score: 121

Sequence: 1 VNLQLLARCVLSNQAYSAYILK 24

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh
-O=/cgn2_1/USPTO.spool/US10036492/runat_02022004_085034_13476/app_query.fasta_1.199
-DB=Issued Patents NA -SUFFIX=rni -MINMATCH=0.1 -LOEFL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=b1csum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10036492 @CGN 1 1 56 @runat_02022004_085034_13476 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	48.8	2472	1	US-08-425-299A-2 Sequence 2, Appli
2	53	43.8	4011	4	US-09-134-001C-2460 Sequence 2460, Ap
3	53	43.8	12886	4	US-09-453-702B-14 Sequence 14, Appl
4	49	40.5	1098	4	US-09-107-532A-2936 Sequence 2936, Ap
5	48	39.7	897	4	US-09-328-352-812 Sequence 812, Ap
6	48	39.7	2999	2	US-09-014-969-8 Sequence 8, Appli
7	48	39.7	90541	4	US-09-759-359A-3 Sequence 3, Appli
8	47	38.8	1308	4	US-09-526-993-4 Sequence 4, Appli
9	47	38.8	2167	4	US-09-526-993-3 Sequence 3, Appli
10	47	38.8	2636	1	US-08-253-785-1 Sequence 1, Appli
11	47	38.8	3224	4	US-09-526-993-2 Sequence 2, Appli
12	47	38.8	6070	4	US-09-526-993-1 Sequence 1, Appli

C 13	47	38.8	6157	4	US-09-526-993-10 Sequence 10, Appli
C 14	47	38.8	6202	4	US-09-526-993-8 Sequence 8, Appli
C 15	46	38.0	642	4	US-09-134-001C-2523 Sequence 2523, Ap
C 16	46	38.0	1269	4	US-09-601-198-178 Sequence 178, App
C 17	46	38.0	1376	4	US-09-620-312D-76 Sequence 76, Appl
C 18	46	38.0	1499	4	US-09-484-970B-157 Sequence 157, App
C 19	46	38.0	1944	4	US-09-398-395A-31 Sequence 31, Appl
C 20	46	38.0	1944	4	US-09-887-586A-31 Sequence 31, Appl
C 21	46	38.0	1944	4	US-09-895-752-31 Sequence 31, Appl
C 22	46	38.0	1944	4	US-09-903-012B-31 Sequence 31, Appl
C 23	46	38.0	1944	4	US-09-484-970B-168 Sequence 168, App
C 24	46	38.0	10684	3	US-08-618-100B-3 Sequence 3, Appli
C 25	45	37.2	137	1	US-08-299-498A-20 Sequence 20, Appl
C 26	45	37.2	137	5	PCT-US95-10813-20 Sequence 944, App
C 27	45	37.2	1305	4	US-09-328-352-944 Sequence 1824, Ap
C 28	45	37.2	1497	4	US-09-134-001C-1824 Sequence 7, Appli
C 29	45	37.2	2454	2	US-08-359-696-3 Sequence 7, Appli
C 30	45	37.2	2580	4	US-09-201-936-7 Sequence 7, Appli
C 31	45	37.2	3036	4	US-09-548-938A-3 Sequence 3, Appli
C 32	45	37.2	3532	2	US-09-205-204-1 Sequence 1, Appli
C 33	45	37.2	3732	3	US-09-212-971-7 Sequence 7, Appli
C 34	45	37.2	3732	3	US-08-800-929A-7 Sequence 7, Appli
C 35	45	37.2	3732	4	US-09-617-053A-7 Sequence 7, Appli
C 36	45	37.2	116592	4	US-09-818-512-3 Sequence 3, Appli
C 37	45	37.2	1830121	4	US-09-557-884-1 Sequence 1, Appli
C 38	45	37.2	1830121	4	US-09-643-990A-1 Sequence 1, Appli
C 39	45	37.2	1830121	4	US-08-998-416-225 Sequence 225, App
C 40	44.5	36.8	533	3	US-09-191-468-54 Sequence 54, Appl
C 41	44.5	36.8	615	4	US-09-191-468-58 Sequence 58, Appl
C 42	44.5	36.8	615	4	US-09-191-468-60 Sequence 60, Appl
C 43	44.5	36.8	615	4	US-09-191-468-62 Sequence 62, Appl
C 44	44.5	36.8	615	4	US-09-191-468-64 Sequence 64, Appl
C 45	44.5	36.8	615	4	US-09-191-468-64 Sequence 64, Appl

ALIGNMENTS

RESULT 1

US-08-425-299A-2
; Sequence 2, Application US/08425299A
; Patent No. 5726025
; GENERAL INFORMATION:
; APPLICANT: Kirschner, Marc W.
; APPLICANT: King, Randall W.
; APPLICANT: Peters, Jean-Michael
; TITLE OF INVENTION: Assay and Reagents for Detecting Inhibitors
; TITLE OF INVENTION: of Ubiquitin-Dependent Degradation of
; TITLE OF INVENTION: Cell Cycle Regulatory Proteins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,299A
; FILING DATE: 20-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-014
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2472 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2469
US-08-425-299A-2
Alignment Scores:
Pred. No.: 2472
Score: 59.00
Percent Similarity: 68.18%
Best Local Similarity: 54.55%
Query Match: 48.76%
DB: 1
Gaps: 0

US-10-036-492-6 (1-24) x US-08-425-299A-2 (1-2472)
Qy 3 LeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyrIle 22
Db 121 TGTGTTTACTGCAACCTGTTATTACCGCTCAGGAAGGCATATAAAGCATATAGACTC 180

Qy 23 LeuLys 24
Db 181 TTGAAA 186

RESULT 2
US-09-134-001C-2460/c
; Sequence 2460, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2460
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2460
Alignment Scores:
Pred. No.: 8.72
Score: 53.00
Percent Similarity: 69.57%
Best Local Similarity: 47.83%
Query Match: 43.80%
DB: 4
Gaps: 0

US-10-036-492-6 (1-24) x US-09-134-001C-2460 (1-1011)
Qy 1 ValAsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 913 GTAGACTGCTCTACTGCTATCCACCTTAACCTTACCTTATTAACCAAT 854

Qy 21 TyrIleLeu 23
Db 853 TACCTGCTT 845

RESULT 3
US-09-453-702B-14
; Sequence 14, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12886
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-453-702B-14
Alignment Scores:
Pred. No.: 193
Score: 53.00
Percent Similarity: 64.29%
Best Local Similarity: 35.71%
Query Match: 43.80%
DB: 4
Gaps: 1

US-10-036-492-6 (1-24) x US-09-453-702B-14 (1-12886)
Qy 1 ValAsnLeuGlnLeu-----AlaATGCTyTyrLeuSerAsnSerGlnAla 16
Db 9431 GTAAATCTGAGTTAATCAACGCTAATGGGGCAATGTTCTCTTAAGGCGGTCAAGTT 9490

Qy 17 TyrSerAlaTyrTyrIleLeuLys 24
Db 9491 AAGAATGCCTTTTACCTGTTGAA 9514

RESULT 4
US-09-107-532A-2936/c
; Sequence 2936, Application US/09107532A
; Patent No. 8583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

```

```
STREET: 100 Beaver Street
CITY: Walham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/107,532A
  FILING DATE: 30-Jun-1998
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 60/085,598
    FILING DATE: 14 May 1998
    APPLICATION NUMBER: 60/051571
    FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
  NAME: Ariniello, Pamela Deneke
  REGISTRATION NUMBER: 40,489
  REFERENCE/DOCKET NUMBER: GTC-012
  TELEPHONE: (781)893-5007
  TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2936:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1098 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: circular
  MOLECULE TYPE: DNA (genomic)
  HYPOTHETICAL: NO
  ANTI-SENSE: NO
  ORIGINAL SOURCE:
    ORGANISM: Enterococcus faecium
  FEATURE:
    NAME/KEY: misc feature
    LOCATION: (B) LOCATION 1...1098
    SEQUENCE DESCRIPTION: SEQ ID NO: 2936:
US-09-107-532A-2936
Alignment Scores:
Pred. No.: 43.3 Length: 1098
Score: 49.00 Matches: 11
Percent Similarity: 77.27% Conservative: 6
Best Local Similarity: 50.00% Mismatches: 3
Query Match: 40.50% Indels: 2
DB: 4 Gaps: 1
US-10-036-492-6 (1-24) x US-09-107-532A-2936 (1-1098)
QY 3 LeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyrIle 22
Db 217 CTTAAACTAATCTCCGCTTGTCCACGTACAGGCA-----TCCCACTATATT 164
QY 23 LeuIlys 24
Db 163 GTGCAG 158
RESULT 5
US-09-328-352-812
Sequence 812, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
  APPLICANT: Gary L. Breton et al.
  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER.
  FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
  CURRENT APPLICATION NUMBER: US/09/328,352
  CURRENT FILING DATE: 1999-06-04
  NUMBER OF SEQ ID NOS: 8252
  SEQ ID NO 812
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-812
Alignment Scores:
Pred. No.: 49.3 Length: 897
Score: 48.00 Matches: 8
Percent Similarity: 75.00% Conservative: 1
Best Local Similarity: 66.67% Mismatches: 3
Query Match: 39.67% Indels: 0
DB: 4 Gaps: 0
US-10-036-492-6 (1-24) x US-09-328-352-812 (1-897)
QY 9 CysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 354 TGCTACGCCGGAACACGTAAAGCATATTCAGGATAC 389
RESULT 6
US-09-014-969-8/c
Sequence 8, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
  APPLICANT: Jacobs, Kenneth
  APPLICANT: McCoy, John M.
  APPLICANT: Lavallee, Edward R.
  APPLICANT: Racie, Lisa A.
  APPLICANT: Merberg, David
  APPLICANT: Treacy, Maurice
  APPLICANT: Spaulding, Vikki
  APPLICANT: Agostino, Michael J.
  TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
  TITLE OF INVENTION: ENCODING THEM
  NUMBER OF SEQUENCES: 32
  CORRESPONDENCE ADDRESS:
    STREET: 87 Bridgepark Drive
    CITY: Cambridge
    STATE: MA
    COUNTRY: U.S.A.
    ZIP: 02140
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patentn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/014,969
    FILING DATE:
  CLASSIFICATION:
  ATTORNEY/AGENT INFORMATION:
    NAME: Sprunger, Suzanne A.
    REGISTRATION NUMBER: 41,323
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (617) 498-8284
    TELEFAX: (617) 876-5851
  INFORMATION FOR SEQ ID NO: 8:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 2999 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
      MOLECULE TYPE: cDNA
US-09-014-969-8
Alignment Scores:
Pred. No.: 214 Length: 2999
Score: 48.00 Matches: 13
Percent Similarity: 45.45% Conservative: 2
Best Local Similarity: 39.39% Mismatches: 8
Query Match: 39.67% Indels: 10
DB: 2 Gaps: 1
```

```
US-10-036-492-6 (1-24) x US-09-014-969-8 (1-2999)
Qy 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsn-----13
Db 2695 GTCAATCTGCAATGTTGACACAATGTTTACACTGTAAATTTCTGTACAAATTAATGTAT 2636
Qy 14 -----SerGlnAlaTyrSerAlaTyrTyrIleLeu 23
Db 2635 ACTTAGAGATACAGGATAACATTTCTACTATATTTTA 2597

RESULT 7
US-09-759-359A-3
; Sequence 3, Application US/09759359A
; Patent No. 6492153
; GENERAL INFORMATION:
; APPLICANT: ABEU-THREIDEH, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/09/759,359A
; CURRENT FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 90541
; TYPE: DNA
; ORGANISM: Human
US-09-759-359A-3

Alignment Scores:
Pred. No.: 1.34e+04 Length: 90541
Score: 48.00 Matches: 7
Percent Similarity: 87.50% Conservative: 7
Best Local Similarity: 43.75% Mismatches: 2
Query Match: 39.67% Indels: 0
DB: 4 Gaps: 0

US-10-036-492-6 (1-24) x US-09-759-359A-3 (1-90541)
Qy 9 CysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyrIleLeuLys 24
Db 73862 TGTTCCTGCGCCATTCAGCAGCTTTTCTCAGTGTATTACTTAGG 73909

RESULT 8
US-09-526-993-4/c
; Sequence 4, Application US/09526993
; Patent No. 6465715
; GENERAL INFORMATION:
; APPLICANT: Zwaal, Richard
; APPLICANT: Asaert, Wouter
; APPLICANT: Roelens, Ingele
; APPLICANT: Bogaert, Thierry
; TITLE OF INVENTION: EXPRESSION OF DNA OR PROTEINS IN C. ELEGANS
; FILE REFERENCE: B0192/7012/ERG/KA
; CURRENT APPLICATION NUMBER: US/09/526,993
; CURRENT FILING DATE: 2000-03-16
; EARLIER FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1308
; TYPE: DNA
; ORGANISM: Caenorhabditis Elegans
US-09-526-993-4

Alignment Scores:
Pred. No.: 114 Length: 1308
Score: 47.00 Matches: 11
Percent Similarity: 58.33% Conservative: 3
Best Local Similarity: 45.83% Mismatches: 10

US-10-036-492-6 (1-24) x US-09-526-993-4 (1-1308)
Qy 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 356 ATAAACTTTATTTTATTCGAACGATGATTAATAATTAATCGCTTTTTCAGCGTTT 297
Qy 21 TyrIleLeuLys 24
Db 296 TTAATTTTAAAA 285

RESULT 9
US-09-526-993-3/c
; Sequence 3, Application US/09526993
; Patent No. 6465715
; GENERAL INFORMATION:
; APPLICANT: Zwaal, Richard
; APPLICANT: Asaert, Wouter
; APPLICANT: Roelens, Ingele
; APPLICANT: Bogaert, Thierry
; TITLE OF INVENTION: EXPRESSION OF DNA OR PROTEINS IN C. ELEGANS
; FILE REFERENCE: B0192/7012/ERG/KA
; CURRENT APPLICATION NUMBER: US/09/526,993
; CURRENT FILING DATE: 2000-03-16
; EARLIER FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2167
; TYPE: DNA
; ORGANISM: Caenorhabditis Elegans
US-09-526-993-3

Alignment Scores:
Pred. No.: 210 Length: 2167
Score: 47.00 Matches: 11
Percent Similarity: 58.33% Conservative: 3
Best Local Similarity: 45.83% Mismatches: 10
Query Match: 38.84% Indels: 0
DB: 4 Gaps: 0

US-10-036-492-6 (1-24) x US-09-526-993-3 (1-2167)
Qy 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 356 ATAAACTTTATTTTATTCGAACGATGATTAATAATTAATCGCTTTTTCAGCGTTT 297
Qy 21 TyrIleLeuLys 24
Db 296 TTAATTTTAAAA 285

RESULT 10
US-08-253-785-1/c
; Sequence 1, Application US/08253785
; Patent No. 563363
; GENERAL INFORMATION:
; APPLICANT: Colbert, James T.
; APPLICANT: Held, Bruce M.
; APPLICANT: Wurtele, Eve S.
; APPLICANT: Dietrich, Paul S.
; TITLE OF INVENTION: ROOT PREFERENTIAL PROMOTER
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sandoz Agro, Inc.
; STREET: 975 California Avenue
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM: -
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/253,785
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Marcus-Wyner, Lynn
REGISTRATION NUMBER: 34,869
REFERENCE/DOCKET NUMBER: 135-1089
TELEPHONE: 415/354-3588
TELEFAX: 415/857-1125
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2636 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-253-785-1

Alignment Scores:
Pred. No.: 266 Length: 2636
Score: 47.00 Matches: 8
Percent Similarity: 57.89% Conservative: 3
Best Local Similarity: 42.11% Mismatches: 8
Query Match: 38.84% Indels: 0
DB: 1 Gaps: 0

US-10-036-492-6 (1-24) x US-08-253-785-1 (1-2636)

Qy 5 LeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyrLeuLeu 23
Db 564 CTTTTCGCGAGTGTACACTCGACAAAGTGACCAATATATTTTATTTGCTT 508

RESULT 11

US-09-526-993-2/c
Sequence 2, Application US/09526993
Patent No. 6465715
GENERAL INFORMATION:
APPLICANT: Zwaal, Richard
APPLICANT: Assaert, Wouter
APPLICANT: Roelens, Ingele
APPLICANT: Bogaert, Thierry
TITLE OF INVENTION: EXPRESSION OF DNA OR PROTEINS IN C. ELEGANS
FILE REFERENCE: B0192/7012/ERG/KA
CURRENT APPLICATION NUMBER: US/09/526,993
CURRENT FILING DATE: 2000-03-16
EARLIER APPLICATION NUMBER: U.K. 9906018.8
EARLIER FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 3224
TYPE: DNA
ORGANISM: Caenorhabditis Elegans
US-09-526-993-2

Alignment Scores:
Pred. No.: 340 Length: 3224
Score: 47.00 Matches: 11
Percent Similarity: 58.33% Conservative: 3
Best Local Similarity: 45.83% Mismatches: 10
Query Match: 38.84% Indels: 0
DB: 4 Gaps: 0

US-10-036-492-6 (1-24) x US-09-526-993-2 (1-3224)

Qy 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 1413 ATAACTTTATTTTATTCGAACGATGATTTAAATAATTTAATCGCTTTTTCAGCGTTT 1354

Qy 21 TyrIleLeuLys 24
Db 1353 TTAATTTTAAAA 1342

RESULT 12

US-09-526-993-1/c
Sequence 1, Application US/09526993
Patent No. 6465715
GENERAL INFORMATION:
APPLICANT: Zwaal, Richard
APPLICANT: Assaert, Wouter
APPLICANT: Roelens, Ingele
APPLICANT: Bogaert, Thierry
TITLE OF INVENTION: EXPRESSION OF DNA OR PROTEINS IN C. ELEGANS
FILE REFERENCE: B0192/7012/ERG/KA
CURRENT APPLICATION NUMBER: US/09/526,993
CURRENT FILING DATE: 2000-03-16
EARLIER APPLICATION NUMBER: U.K. 9906018.8
EARLIER FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 6070
TYPE: DNA
ORGANISM: Caenorhabditis Elegans
US-09-526-993-1

Alignment Scores:
Pred. No.: 735 Length: 6070
Score: 47.00 Matches: 11
Percent Similarity: 58.33% Conservative: 3
Best Local Similarity: 45.83% Mismatches: 10
Query Match: 38.84% Indels: 0
DB: 4 Gaps: 0

US-10-036-492-6 (1-24) x US-09-526-993-1 (1-6070)

Qy 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 4259 ATAACTTTATTTTATTCGAACGATGATTTAAATAATTTAATCGCTTTTTCAGCGTTT 4200

Qy 21 TyrIleLeuLys 24
Db 4199 TTAATTTTAAAA 4188

RESULT 13

US-09-526-993-10/c
Sequence 10, Application US/09526993
Patent No. 6465715
GENERAL INFORMATION:
APPLICANT: Zwaal, Richard
APPLICANT: Assaert, Wouter
APPLICANT: Roelens, Ingele
APPLICANT: Bogaert, Thierry
TITLE OF INVENTION: EXPRESSION OF DNA OR PROTEINS IN C. ELEGANS
FILE REFERENCE: B0192/7012/ERG/KA
CURRENT APPLICATION NUMBER: US/09/526,993
CURRENT FILING DATE: 2000-03-16
EARLIER APPLICATION NUMBER: U.K. 9906018.8
EARLIER FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 6157
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: plasmid pGF2013
US-09-526-993-10

Alignment Scores:
Pred. No.: 747 Length: 6157

Score: 47.00 Matches: 11
Percent Similarity: 58.33% Conservative: 3
Best Local Similarity: 45.83% Mismatches: 10
Query Match: 38.84% Indels: 0
DB: 4 Gaps: 0

US-10-036-492-6 (1-24) x US-09-526-993-10 (1-6157)

QY 1 ValAsnLeuGlnIleuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
DB 2446 ATAACTTTATTTATTCGACGATGCAATTAATAATTTAATCGCTTTTCACGGTTT 2387
QY 21 TyrIleLeuLys 24
DB 2386 TTAATTTTAAAA 2375

RESULT 14

US-09-526-993-8/c
; Sequence 8, Application US/09526993
; Patent No. 6465715
; GENERAL INFORMATION:
; APPLICANT: Zwaal, Richard
; APPLICANT: Asaert, Wouter
; APPLICANT: Roelens, Ingele
; APPLICANT: Bogaert, Thierry
; TITLE OF INVENTION: EXPRESSION OF DNA OR PROTEINS IN C. ELEGANS
; FILE REFERENCE: B0192/7012/ERG/KA
; CURRENT APPLICATION NUMBER: US/09/526,993
; CURRENT FILING DATE: 2000-03-16
; EARLIER APPLICATION NUMBER: U.K. 9906018.8
; EARLIER FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 6202
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid pGF2006
US-09-526-993-8

Alignment Scores:

Pred. No.: 754 Length: 6202
Score: 47.00 Matches: 11
Percent Similarity: 58.33% Conservative: 3
Best Local Similarity: 45.83% Mismatches: 10
Query Match: 38.84% Indels: 0
DB: 4 Gaps: 0

US-10-036-492-6 (1-24) x US-09-526-993-8 (1-6202)

QY 1 ValAsnLeuGlnIleuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
DB 361 ATAACTTTATTTATTCGACGATGCAATTAATAATTTAATCGCTTTTCACGGTTT 302
QY 21 TyrIleLeuLys 24
DB 301 TTAATTTTAAAA 290

RESULT 15

US-09-134-001C-2523
; Sequence 2523, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GFC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2523
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2523

Alignment Scores:
Pred. No.: 69.6 Length: 642
Score: 46.00 Matches: 9
Percent Similarity: 69.23% Conservative: 0
Best Local Similarity: 69.23% Mismatches: 4
Query Match: 38.02% Indels: 0
DB: 4 Gaps: 0

US-10-036-492-6 (1-24) x US-09-134-001C-2523 (1-642)

QY 8 ArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
DB 265 CGATGTTACAACTCTAACTATGATTTTACAGCGCATAT 303

Search completed: February 2, 2004, 17:03:13
Job time : 170 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 2, 2004, 13:02:44 ; Search time 1605 Seconds
(without alignments)
40.365 Million cell updates/sec

Title: US-10-036-492-6

Perfect score: 121

Sequence: 1 VNLQLARCYLSNSQAYSAYILK 24

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-WDEL=frame+ps2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_epool/US10036492/runat 02022004 085033 13450/app_query.fasta_1.199
-DB=N Geneseq 19Jun03 -QMT=fastat -SUFFIX=ring -MNMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10036492 @CGN 1.1 312 @runat 02022004 085033 13450 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=0.5 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N Geneseq 19Jun03.*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	100.0	2401	22	Arabidopsis CDC27A
2	121	100.0	2434	22	Arabidopsis CDC27A
3	91	75.2	486	24	DNA encoding cell
4	91	75.2	498	24	DNA encoding cell
5	91	75.2	1557	24	DNA encoding funct
6	91	75.2	2115	24	DNA encoding funct
7	91	75.2	2169	24	DNA encoding funct
8	91	75.2	2220	22	Arabidopsis CDC27B
9	91	75.2	2235	24	DNA encoding cell
10	91	75.2	2235	24	DNA encoding funct
11	91	75.2	2235	24	DNA encoding funct
12	91	75.2	2235	24	DNA encoding funct
13	91	75.2	2235	24	DNA encoding funct
14	91	75.2	2316	21	Arabidopsis thalia
15	91	75.2	2480	24	DNA encoding cell
16	91	75.2	4577	24	Genomic DNA encodi
17	91	75.2	4577	24	DNA encoding funct
18	91	75.2	4577	24	DNA encoding funct
19	91	75.2	4577	24	DNA encoding funct
20	91	75.2	4577	24	DNA encoding funct
21	91	75.2	4577	24	DNA encoding funct
22	91	75.2	4577	24	DNA encoding funct
23	91	75.2	4577	24	DNA encoding funct
24	59	48.8	314	21	Human secreted pro
25	59	48.8	486	14	Human brain expres
26	59	48.8	2472	17	Human CDC27 codin
27	59	48.8	2472	17	Human CDC27 codin
28	59	48.8	2747	23	DNA encoding novel
29	59	48.8	3320	16	H-NUC retinoblasto
30	53	43.8	411	22	S. epidermidis ope
31	53	43.8	990	22	S. epidermidis ope
32	53	43.8	1011	24	Staphylococcus epi
33	53	43.8	3330	22	S. epidermidis gen
34	53	43.8	3532	22	S. epidermidis gen
35	52	43.0	1082138	21	Arabidopsis thalia
36	52	43.0	1272	25	Pathogen specific
37	52	43.0	1987	18	Staphylococcus aur
38	52	43.0	4354	23	Drosophila melanog
39	52	43.0	5519	23	Drosophila melanog
40	52	43.0	26270	21	Haemorrhagic enter
41	51	42.1	383	25	Human GDP-mannose
42	51	42.1	412	24	Human ovarian canc
43	51	42.1	502	24	Human colon cancer
44	51	42.1	592	24	Human colon cancer
45	51	42.1	1415	24	Human cDNA differe

ALIGNMENTS

RESULT 1

AAF56506
ID AAF56506 standard; cDNA; 2401 BP.

XX AAF56506;

XX AAF56506;

XX 18-APR-2001 (first entry)

XX Arabidopsis CDC27A2 coding sequence.

XX Cell cycle regulation; DNA replication; CDC7; CDC27A1; CDC27A2;
XX CDC27B; nematode resistance; endoreduplication; sterility;
XX polyploidy; ss.

XX Arabidopsis thaliana.

XX WO200102430-A2.

```

XX PD 11-JAN-2001.
XX PF
XX PR 05-JUL-2000; 2000WO-EP06401.
XX PR 05-JUL-1999; 99EP-0202214.
XX PA (CROP-) CROPDISEIGN NV.
XX PA (UYRI-) UNIV RIO DE JANEIRO.
XX PI Hemerly AS, Ferreira PCG, Rombauts S;
XX PR WPI; 2001-123101/13.
XX PR Partially purified plant CDC27 or CDC7 protein homolog, useful for
XX PT modulating DNA replication and for producing transgenic plants
XX PS Claim 13; Page 85-86; 86pp; English.
XX CC The present invention provides the protein and coding sequences of
XX CC several Arabidopsis thaliana proteins which are involved in DNA
XX CC replication and the regulation of the cell cycle. These include CDC7,
XX CC CDC27A1, CDC27A2 and CDC27B. They are useful in the production of
XX CC transgenic and mutant plants, as the mutations in the gene cause
XX CC proteins to confer nematode resistance, sterility and polyploidy on
XX CC plants and also lead to endoreduplication.
XX SQ Sequence 2401 BP; 747 A; 466 C; 514 G; 674 T; 0 other;

Alignment Scores:
Pred. No.: 2,88e-09 Length: 2401
Score: 121.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-036-492-6 (1-24) x AAF56506 (1-2401)

QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 109 GTGAACCTGCAATTGTAGCCAGGTGTACTTGGTAAACAGTCAAGCTTATAGTCATAT 168

QY 21 TyriLeuLeuLys 24
Db 169 TATATCCTTAAA 180

RESULT 2
AAF56505
ID AAF56505 standard; cDNA; 2434 BP.
XX AC AAF56505;
XX DT 18-APR-2001 (first entry)
XX DE Arabidopsis CDC27A1 coding sequence.
XX KW Cell cycle regulation; DNA replication; CDC7; CDC27A1; CDC27A2;
XX KW CDC27B; nematode resistance; endoreduplication; sterility;
XX KW polyploidy; ss.
XX OS Arabidopsis thaliana.
XX FN WO200102430-A2.
XX PD 11-JAN-2001.
XX PR 05-JUL-2000; 2000WO-EP06401.
XX PR 05-JUL-1999; 99EP-0202214.
XX PA (CROP-) CROPDISEIGN NV.
XX PA (UYRI-) UNIV RIO DE JANEIRO.

```

```

XX PI Hemerly AS, Ferreira PCG, Rombauts S;
XX DR WPI; 2001-123101/13.
XX PR Partially purified plant CDC27 or CDC7 protein homolog, useful for
XX PT modulating DNA replication and for producing transgenic plants
XX PS Claim 13; Page 76-77; 86pp; English.
XX CC The present invention provides the protein and coding sequences of
XX CC several Arabidopsis thaliana proteins which are involved in DNA
XX CC replication and the regulation of the cell cycle. These include CDC7,
XX CC CDC27A1, CDC27A2 and CDC27B. They are useful in the production of
XX CC transgenic and mutant plants, as the mutations in the gene cause
XX CC proteins to confer nematode resistance, sterility and polyploidy on
XX CC plants and also lead to endoreduplication.
XX SQ Sequence 2434 BP; 755 A; 475 C; 520 G; 684 T; 0 other;

Alignment Scores:
Pred. No.: 2,93e-09 Length: 2434
Score: 121.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-036-492-6 (1-24) x AAF56505 (1-2434)

QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 109 GTGAACCTGCAATTGTAGCCAGGTGTACTTGGTAAACAGTCAAGCTTATAGTCATAT 168

QY 21 TyriLeuLeuLys 24
Db 169 TATATCCTTAAA 180

RESULT 3
ABK93343
ID ABK93343 standard; cDNA; 486 BP.
XX AC ABK93343;
XX DT 23-AUG-2002 (first entry)
XX DE DNA encoding cell cycle regulation protein HOBBIT (HBT) #3.
XX KW Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
XX KW pattern formation; plant meristem development; plant yield; drought;
XX KW seedling emergency; root generation; shade avoidance response;
XX KW vascular strand formation; patterning; parthenocarpic fruit;
XX KW cell cycle regulation; endoreduplication; cell division;
XX KW transgenic plant; plant development; plant morphology; plant physiology;
XX KW plant biochemistry; gene; ss.
XX OS Arabidopsis thaliana.
XX FN WO200238599-A2.
XX PD 16-MAY-2002.
XX PR 13-NOV-2001; 2001WO-EP13116.
XX PR 13-NOV-2000; 2000EP-0870271.
XX PR 30-NOV-2000; 2000US-250402P.
XX PA (UYUT-) RIJKSUNIV UTRECHT.
XX PI Scheres BJG, Billaou I, Folmer SDH;
XX DR WPI; 2002-490065/52.
XX DR P-PSDB; ABG65520.

```

XX OS Arabidopsis thaliana.
 XX PN WO200238599-A2.
 XX PD 16-MAY-2002.
 XX PF 13-NOV-2001; 2001WO-EPI31116.
 XX PR 13-NOV-2000; 2000EP-0870271.
 XX PR 30-NOV-2000; 2000US-250402P.
 XX PA (UYUT-) RIJKSUNIV UTRECHT.
 XX PI Scheres BJG, Biiilou I, Folmer SDH;
 XX WPI; 2002-490065/52.
 XX DR P-PSDB; ABG65521.
 XX PT Use of plant cdc27B for modulating or mimicking auxin-related effects
 XX in plants or plant cells, or for regulation of cell cycle of plant cell
 XX Claim 31; Page 143; 207pp; English.
 XX The invention described the use of a plant cdc27B (also termed HOBBIT
 CC (HBT)) useful for modulating or mimicking auxin-related effects in a
 CC plant or plant cell. The modulation or mimicking of auxin-related effects
 CC results in: altered cell fate and/or altered pattern formation in a plant
 CC or plant cell; alteration in the size and/or number of naturally
 CC occurring plant meristems; modification of the numbers of organs or
 CC tissues, and/or a modification of the rate of organ or tissue emanation
 CC from a plant meristem, and/or a modification of the arrangement of organs
 CC and/or tissues in a plant; increased survival rate under drought conditions;
 CC rate of plants; an enhanced emergency; mimicking of root generation in tissue
 CC cultures; an increased shade avoidance response; altering vascular strand
 CC formation and patterning in a plant; or production of parthenocarpic
 CC fruits. HBT also regulates the cell cycle of the plant cell results in:
 CC modulated endoreduplication in a plant; modulation of the rate of cell
 CC division; alteration in the size of naturally occurring meristems; a
 CC modification in a number of organs or tissues, and/or the modification of
 CC the rate of organ or tissue emanation from a plant meristem and/or
 CC and/or an increased plant yield or enhancement of survival rate of
 CC plants. A nucleic acid encoding HBT is useful for producing transgenic
 CC plants, plant cells or plant tissues and in the regeneration of a plant
 CC from a plant cell. HBT nucleic acids and proteins are useful for
 CC modifying cell fate, pattern formation, plant development, plant
 CC morphology, plant physiology and/or plant biochemistry. This sequence
 CC encodes an Arabidopsis thaliana HOBBIT (HBT) or cdc27B protein described
 CC in the invention.
 XX SQ Sequence 486 BP; 132 A; 108 C; 102 G; 144 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,33e-05 Length: 486
 Score: 91.00 Matches: 18
 Percent Similarity: 87.50% Conservative: 3
 Best Local Similarity: 75.00% Mismatches: 3
 Query Match: 75.21% Indels: 0
 DB: 24 Gaps: 0
 US-10-036-492-6 (1-24) x ABK93343 (1-486)
 QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 Db 106 GTTAATTGCGCTATTAGCCACGACTACCTCGAGAATAATCAAGCTTACAGTGCATAT 165
 QY 21 TyrIleLeuLys 24
 Db 166 CATCTGCTAAG 177
 RESULT 4
 ABK93344
 ID ABK93344 standard; cDNA; 498 BP.
 XX AC ABK93344;
 XX DT 23-AUG-2002 (first entry)
 XX DE DNA encoding cell cycle regulation protein HOBBIT (HBT) #4.
 XX KW plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
 KW pattern formation; plant meristem development; plant yield; drought;
 KW seedling emergency; root generation; shade avoidance response;
 KW vascular strand formation; patterning; parthenocarpic fruit;
 KW cell cycle regulation; endoreduplication; cell division;
 KW transgenic plant; plant development; plant morphology; plant physiology;
 KW plant biochemistry; gene; ss.

XX OS Arabidopsis thaliana.
 XX PN WO200238599-A2.
 XX PD 16-MAY-2002.
 XX PF 13-NOV-2001; 2001WO-EPI31116.
 XX PR 13-NOV-2000; 2000EP-0870271.
 XX PR 30-NOV-2000; 2000US-250402P.
 XX PA (UYUT-) RIJKSUNIV UTRECHT.
 XX PI Scheres BJG, Biiilou I, Folmer SDH;
 XX WPI; 2002-490065/52.
 XX DR P-PSDB; ABG65521.
 XX PT Use of plant cdc27B for modulating or mimicking auxin-related effects
 XX in plants or plant cells, or for regulation of cell cycle of plant cell
 XX Claim 31; Page 144; 207pp; English.
 XX The invention described the use of a plant cdc27B (also termed HOBBIT
 CC (HBT)) useful for modulating or mimicking auxin-related effects in a
 CC plant or plant cell. The modulation or mimicking of auxin-related effects
 CC results in: altered cell fate and/or altered pattern formation in a plant
 CC or plant cell; alteration in the size and/or number of naturally
 CC occurring plant meristems; modification of the numbers of organs or
 CC tissues, and/or a modification of the rate of organ or tissue emanation
 CC from a plant meristem, and/or a modification of the arrangement of organs
 CC and/or tissues in a plant; increased survival rate under drought conditions;
 CC rate of plants; an enhanced emergency; mimicking of root generation in tissue
 CC cultures; an increased shade avoidance response; altering vascular strand
 CC formation and patterning in a plant; or production of parthenocarpic
 CC fruits. HBT also regulates the cell cycle of the plant cell results in:
 CC modulated endoreduplication in a plant; modulation of the rate of cell
 CC division; alteration in the size of naturally occurring meristems; a
 CC modification in a number of organs or tissues, and/or the modification of
 CC the rate of organ or tissue emanation from a plant meristem and/or
 CC and/or an increased plant yield or enhancement of survival rate of
 CC plants. A nucleic acid encoding HBT is useful for producing transgenic
 CC plants, plant cells or plant tissues and in the regeneration of a plant
 CC from a plant cell. HBT nucleic acids and proteins are useful for
 CC modifying cell fate, pattern formation, plant development, plant
 CC morphology, plant physiology and/or plant biochemistry. This sequence
 CC encodes an Arabidopsis thaliana HOBBIT (HBT) or cdc27B protein described
 CC in the invention.
 XX SQ Sequence 498 BP; 136 A; 109 C; 105 G; 148 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,4e-05 Length: 498
 Score: 91.00 Matches: 18
 Percent Similarity: 87.50% Conservative: 3
 Best Local Similarity: 75.00% Mismatches: 3
 Query Match: 75.21% Indels: 0
 DB: 24 Gaps: 0
 US-10-036-492-6 (1-24) x ABK93344 (1-498)
 QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 Db 106 GTTAATTGCGCTATTAGCCACGACTACCTCGAGAATAATCAAGCTTACAGTGCATAT 165
 QY 21 TyrIleLeuLys 24
 Db 166 CATCTGCTAAG 177

RESULT 5
 ABK93357
 ID ABK93357 standard; cDNA; 1557 BP.
 XX
 AC ABK93357;
 XX
 DT 23-AUG-2002 (first entry)
 XX
 DE DNA encoding functionally inactive HOBBIT (HBT) protein #13.
 XX
 KW Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
 KW pattern formation; plant meristem development; plant yield; drought;
 KW seedling emergency; root generation; shade avoidance response;
 KW vascular strand formation; patterning; parthenocarpic fruit;
 KW cell cycle regulation; endoreduplication; cell division;
 KW transgenic plant; plant development; plant morphology; plant physiology;
 KW plant biochemistry; gene; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200238599-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 13-NOV-2001; 2001WO-EPI31116.
 XX
 PR 13-NOV-2000; 2000EP-0870271.
 PR 30-NOV-2000; 2000US-250402P.
 XX
 PA (UYUT-) RIJXSUNIV UTRECHT.
 XX
 PI Scheres BJG, Bllou I, Folmer SDH;
 XX
 WPI; 2002-490065/52.
 XX
 Use of plant cdc27B for modulating or mimicking auxin-related effects
 in plants or plant cells, or for regulation of cell cycle of plant cell
 .
 Claim 36; Page 176-177; 207pp; English.
 XX
 The invention described the use of a plant cdc27B (also termed HOBBIT
 (HBT)) useful for modulating or mimicking auxin-related effects in a
 plant or plant cell. The modulation or mimicking of auxin-related effects
 results in: altered cell fate and/or altered pattern formation in a plant
 or plant cell; alteration in the size and/or number of naturally
 occurring plant meristems; modification of the rate of organ or tissue emanation
 from a plant meristem, and/or a modification of the arrangement of organs
 and/or tissues in a plant; increased plant yield; an increased survival
 rate of plants; an enhanced survival rate under drought conditions;
 increase of seedling emergency; mimicking of root generation in tissue
 cultures; an increased shade avoidance response; altering vascular strand
 formation and patterning in a plant; or production of parthenocarpic
 fruits. HBT also regulates the cell cycle of the plant cell results in:
 modulated endoreduplication in a plant; modulation of sterility in
 plants; increasing the cell cycle or increase of the rate of cell
 division; alteration in the size of naturally occurring meristems; a
 modification in a number of organs or tissues, and/or the modification of
 the rate of organ or tissue emanation from a plant meristem and/or
 modification of the arrangement of the organs and/or tissues in a plant
 and/or an increased plant yield or enhancement of survival rate of
 plants. A nucleic acid encoding HBT is useful for producing transgenic
 plants, plant cells or plant tissues and in the regeneration of a plant
 from a plant cell. HBT nucleic acids and proteins are useful for
 modifying cell fate, pattern formation, plant development, plant
 morphology, plant physiology and/or plant biochemistry. This sequence
 encodes a functionally inactive Arabidopsis thaliana HOBBIT (HBT) or
 cdc27B protein described in the invention.
 XX
 SQ Sequence 1557 BP; 448 A; 339 C; 328 G; 442 T; 0 other;

Alignment Scores:
 Pred. No.: 9.58e-05 Length: 1557
 Score: 91.00 Matches: 18
 Percent Similarity: 87.50% Conservative: 3
 Best Local Similarity: 75.00% Mismatches: 3
 Query Match: 75.21% Indels: 0
 DB: 24 Gaps: 0

US-10-036-492-6 (1-24) x ABK93357 (1-1557)
 QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 DB 106 GTTAATTGAGCTATTAGCCACCACCTACCTGCAGATAATCAAGCTTACAGTGCATAT 165
 QY 21 TyrIleLeuLys 24
 DB 166 CATCTGCTAAAG 177

RESULT 6
 ABK93356
 ID ABK93356 standard; cDNA; 2115 BP.
 XX
 AC ABK93356;
 XX
 DT 23-AUG-2002 (first entry)
 XX
 DE DNA encoding functionally inactive HOBBIT (HBT) protein #12.
 XX
 KW Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
 KW pattern formation; plant meristem development; plant yield; drought;
 KW seedling emergency; root generation; shade avoidance response;
 KW vascular strand formation; patterning; parthenocarpic fruit;
 KW cell cycle regulation; endoreduplication; cell division;
 KW transgenic plant; plant development; plant morphology; plant physiology;
 KW plant biochemistry; gene; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO200238599-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 13-NOV-2001; 2001WO-EPI31116.
 XX
 PR 13-NOV-2000; 2000EP-0870271.
 PR 30-NOV-2000; 2000US-250402P.
 XX
 PA (UYUT-) RIJXSUNIV UTRECHT.
 XX
 PI Scheres BJG, Bllou I, Folmer SDH;
 XX
 WPI; 2002-490065/52.
 XX
 Use of plant cdc27B for modulating or mimicking auxin-related effects
 in plants or plant cells, or for regulation of cell cycle of plant cell
 .
 Claim 36; Page 175-176; 207pp; English.
 XX
 The invention described the use of a plant cdc27B (also termed HOBBIT
 (HBT)) useful for modulating or mimicking auxin-related effects in a
 plant or plant cell. The modulation or mimicking of auxin-related effects
 results in: altered cell fate and/or altered pattern formation in a plant
 or plant cell; alteration in the size and/or number of naturally
 occurring plant meristems; modification of the rate of organ or tissue emanation
 from a plant meristem, and/or a modification of the arrangement of organs
 and/or tissues in a plant; increased plant yield; an increased survival
 rate of plants; an enhanced survival rate under drought conditions;
 increase of seedling emergency; mimicking of root generation in tissue
 cultures; an increased shade avoidance response; altering vascular strand
 formation and patterning in a plant; or production of parthenocarpic
 fruits. HBT also regulates the cell cycle of the plant cell results in:
 modulated endoreduplication in a plant; modulation of sterility in
 plants; increasing the cell cycle or increase of the rate of cell
 division; alteration in the size of naturally occurring meristems; a
 modification in a number of organs or tissues, and/or the modification of
 the rate of organ or tissue emanation from a plant meristem and/or
 modification of the arrangement of the organs and/or tissues in a plant
 and/or an increased plant yield or enhancement of survival rate of
 plants. A nucleic acid encoding HBT is useful for producing transgenic
 plants, plant cells or plant tissues and in the regeneration of a plant
 from a plant cell. HBT nucleic acids and proteins are useful for
 modifying cell fate, pattern formation, plant development, plant
 morphology, plant physiology and/or plant biochemistry. This sequence
 encodes a functionally inactive Arabidopsis thaliana HOBBIT (HBT) or
 cdc27B protein described in the invention.
 XX
 SQ Sequence 1557 BP; 448 A; 339 C; 328 G; 442 T; 0 other;

CC modulated endoreduplication in a plant; modulation of sterility in
 CC plants; increasing the cell cycle or increase of the rate of cell
 CC division; alteration in the size of naturally occurring meristems; a
 CC modification in a number of organs or tissues, and/or the modification of
 CC the rate of organ or tissue emanation from a plant meristem and/or
 CC modification of the arrangement of the organs and/or tissues in a plant
 CC and/or an increased plant yield or enhancement of survival rate of
 CC plants. A nucleic acid encoding HBT is useful for producing transgenic
 CC plants, plant cells or plant tissues and in the regeneration of a plant
 CC from a plant cell. HBT nucleic acids and proteins are useful for
 CC modifying cell fate, pattern formation, plant development, plant
 CC morphology, plant physiology and/or plant biochemistry. This sequence
 CC encodes a functionally inactive Arabidopsis thaliana HOBBIT (HBT) or
 CC cdc27B protein described in the invention.

XX
 SQ Sequence 2115 BP; 634 A; 450 C; 458 G; 573 T; 0 other;

Alignment Scores:
 Pred. No.: 0.000139 Length: 2115
 Score: 91.00 Matches: 18
 Percent Similarity: 87.50% Conservative: 3
 Best Local Similarity: 75.00% Mismatches: 3
 Query Match: 75.21% Indels: 0
 DB: 24 Gaps: 0

US-10-036-492-6 (1-24) x ABK93356 (1-2115)

QY 1 ValAsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 Db 106 GTTAAATTGACGCTATTAGCCACGCTACCTGCAGATAATCAAGCTTACAGTGCATAT 165
 QY 21 TyrIleLeuLys 24
 Db 166 CATCTGCTAAAG 177

RESULT 7

ID ABK93355 standard; cDNA; 2169 BP.

AC ABK93355;

DT 23-AUG-2002 (first entry)

DE DNA encoding functionally inactive HOBBIT (HBT) protein #11.

KW Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
 KW pattern formation; plant meristem development; plant yield; drought;
 KW seedling emergency; root generation; shade avoidance response;
 KW vascular strand formation; patterning; parthenocarpic fruit;
 KW cell cycle regulation; endoreduplication; cell division;
 KW transgenic plant; plant development; plant morphology; plant physiology;
 KW plant biochemistry; gene; ss.

OS Arabidopsis thaliana.

PN WO200238599-A2.

PD 16-MAY-2002.

PF 13-NOV-2001; 2001WO-EP13116.

PR 13-NOV-2000; 2000EP-0870271.

PR 30-NOV-2000; 2000US-250402P.

PA (UYUT-) RIJKSUNIV UTRECHT.

PI Scheres BJG, Blilou I, Folmer SDH;

XX WPI; 2002-490065/52.

XX Use of plant cdc27B for modulating or mimicking auxin-related effects
 PT in plants or plant cells, or for regulation of cell cycle of plant cell

XX

XX Claim 36; Page 174-175; 207pp; English.

CC The invention described the use of a plant cdc27B (also termed HOBBIT
 (HBT)) useful for modulating or mimicking auxin-related effects in a
 CC plant or plant cell. The modulation or mimicking of auxin-related effects
 CC results in: altered cell fate and/or altered pattern formation in a plant
 CC or plant cell; alteration in the size and/or number of naturally
 CC occurring plant meristems; modification of the numbers of organs or
 CC tissues, and/or a modification of the rate of organ or tissue emanation
 CC from a plant meristem, and/or a modification of the arrangement of organs
 CC and/or tissues in a plant; increased plant yield; an increased survival
 CC rate of plants; an enhanced survival rate under drought conditions;
 CC increase of seedling emergency; mimicking of root generation in tissue
 CC cultures; an increased shade avoidance response; altering vascular strand
 CC formation and patterning in a plant; or production of parthenocarpic
 CC fruits. HBT also regulates the cell cycle of the plant cell results in:
 CC modulated endoreduplication in a plant; modulation of sterility in
 CC plants; increasing the cell cycle or increase of the rate of cell
 CC division; alteration in the size of naturally occurring meristems; a
 CC modification in a number of organs or tissues, and/or the modification of
 CC the rate of organ or tissue emanation from a plant meristem and/or
 CC modification of the arrangement of the organs and/or tissues in a plant
 CC and/or an increased plant yield or enhancement of survival rate of
 CC plants. A nucleic acid encoding HBT is useful for producing transgenic
 CC plants, plant cells or plant tissues and in the regeneration of a plant
 CC from a plant cell. HBT nucleic acids and proteins are useful for
 CC modifying cell fate, pattern formation, plant development, plant
 CC morphology, plant physiology and/or plant biochemistry. This sequence
 CC encodes a functionally inactive Arabidopsis thaliana HOBBIT (HBT) or
 CC cdc27B protein described in the invention.

XX
 SQ Sequence 2169 BP; 643 A; 462 C; 470 G; 594 T; 0 other;

Alignment Scores:

Pred. No.: 0.000143 Length: 2169
 Score: 91.00 Matches: 18
 Percent Similarity: 87.50% Conservative: 3
 Best Local Similarity: 75.00% Mismatches: 3
 Query Match: 75.21% Indels: 0
 DB: 24 Gaps: 0

US-10-036-492-6 (1-24) x ABK93355 (1-2169)

QY 1 ValAsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 Db 106 GTTAAATTGACGCTATTAGCCACGCTACCTGCAGATAATCAAGCTTACAGTGCATAT 165
 QY 21 TyrIleLeuLys 24
 Db 166 CATCTGCTAAAG 177

RESULT 8

AAF56507

ID AAF56507 standard; cDNA; 2220 BP.

XX

AC AAF56507;

DT 18-APR-2001 (first entry)

DE Arabidopsis cdc27B coding sequence.

XX Cell cycle regulation; DNA replication; CDC7; CDC27A1; CDC27A2;
 KW CDC27B; nematode resistance; endoreduplication; sterility;
 KW polyploidy; ss.

OS Arabidopsis thaliana.

PN WO200102430-A2.

XX

PD 11-JAN-2001.

XX 05-JUL-2000; 2000WO-EP06401.

```

XX PR 05-JUL-1999; 99EP-0202214.
XX PA (CROP-) CROPDISEGN NV.
XX PA (UYRI-) UNIV RIO DE JANEIRO.
XX PI Hemerly AS, Ferreira PCG, Rombauts S;
XX PS WPI; 2001-123101/13.
XX PT Partially purified plant CDC27 or CDC7 protein homolog, useful for
XX PT modulating DNA replication and for producing transgenic plants
XX PS Claim 13; Page 86; 86pp; English.
XX CC The present invention provides the protein and coding sequences of
XX CC several Arabidopsis thaliana proteins which are involved in DNA
XX CC replication and the regulation of the cell cycle. These include CDC7,
XX CC CDC27A1, CDC27A2 and CDC27B. They are useful in the production of
XX CC transgenic and mutant plants, as the mutations in the gene cause
XX CC proteins to confer nematode resistance, sterility and polyploidy on
XX CC plants and also lead to endoreduplication.
XX SQ Sequence 2220 BP; 658 A; 474 C; 480 G; 608 T; 0 other;

Alignment Scores:
Pred. NO.: 0.000147 Length: 2220
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 22 Gaps: 0

US-10-036-492-6 (1-24) x AAF56507 (1-2220)
QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 106 GTTAAATTGGAGCTATTAGCCACCAGCTACTCTGCAGATATCAAGCTTACAGTGCATAT 165

QY 21 TyrIleLeuLys 24
Db 166 CATCTGCTAAAG 177

RESULT 9
ABK93342
ID ABK93342 standard; cDNA; 2235 BP.
XX AC ABK93342;
XX XX 23-AUG-2002 (first entry)
XX DE DNA encoding cell cycle regulation protein HOBBIT (HBT) #2.
XX KW Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
XX KW pattern formation; plant meristem development; plant yield; drought;
XX KW seedling emergency; root generation; shade avoidance response;
XX KW vascular strand formation; patterning; parthenocarpic fruit;
XX KW cell cycle regulation; endoreduplication; cell division;
XX KW transgenic plant; plant development; plant morphology; plant physiology;
XX KW plant biochemistry; gene; ss.
XX OS Arabidopsis thaliana.
XX PN WO200238599-A2.
XX PD 16-MAY-2002.
XX XX 13-NOV-2001; 2001WO-EF13116.
XX PF 13-NOV-2000; 2000EP-0870271.
XX PR 30-NOV-2000; 2000US-250402P.
XX XX (UYUT-) RIJKSUNIV UTRECHT.
XX PA

```

```

XX PI Scheres BJG, Bllilou I, Folmer SDH;
XX DR WPI; 2002-490065/52.
XX PT Use of plant cdc27B for modulating or mimicking auxin-related effects
XX PT in plants or plant cells, or for regulation of cell cycle of plant cell
XX PS Claim 31; Page 142-143; 207pp; English.
XX CC The invention described the use of a plant cdc27B (also termed HOBBIT
XX CC (HBT)) useful for modulating or mimicking auxin-related effects in a
XX CC plant or plant cell. The modulation or mimicking of auxin-related effects
XX CC results in: altered cell fate and/or altered pattern formation in a plant
XX CC or plant cell; alteration in the size and/or number of naturally
XX CC occurring plant meristems; modification of the numbers of organs or
XX CC tissues, and/or a modification of the rate of organ or tissue emanation
XX CC from a plant meristem, and/or a modification of the arrangement of organs
XX CC and/or tissues in a plant; increased plant yield; an increased survival
XX CC rate of plants; an enhanced survival rate under drought conditions;
XX CC increase of seedling emergency; mimicking of root generation in tissue
XX CC cultures; an increased shade avoidance response; altering vascular strand
XX CC formation and patterning in a plant; or production of parthenocarpic
XX CC fruits. HBT also regulates the cell cycle of the plant cell results in:
XX CC modulated endoreduplication in a plant; modulation of sterility in
XX CC plants; increasing the cell cycle or increase of the rate of cell
XX CC division; alteration in the size of naturally occurring meristems; a
XX CC modification in a number of organs or tissues, and/or the modification of
XX CC the rate of organ or tissue emanation from a plant meristem and/or
XX CC modification of the arrangement of the organs and/or tissues in a plant
XX CC and/or an increased plant yield or enhancement of survival rate of
XX CC plants. A nucleic acid encoding HBT is useful for producing transgenic
XX CC plants, plant cells or plant tissues and in the regeneration of a plant
XX CC from a plant cell. HBT nucleic acids and proteins are useful for
XX CC modifying cell fate, pattern formation, plant development, plant
XX CC morphology, plant physiology and/or plant biochemistry. This sequence
XX CC encodes an Arabidopsis thaliana HOBBIT (HBT) or cdc27B protein described
XX CC in the invention.
XX SQ Sequence 2235 BP; 664 A; 475 C; 484 G; 612 T; 0 other;

Alignment Scores:
Pred. NO.: 0.000149 Length: 2235
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 24 Gaps: 0

US-10-036-492-6 (1-24) x ABK93342 (1-2235)
QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 106 GTTAAATTGGAGCTATTAGCCACCAGCTACTCTGCAGATATCAAGCTTACAGTGCATAT 165

QY 21 TyrIleLeuLys 24
Db 166 CATCTGCTAAAG 177

RESULT 10
ABK93352
ID ABK93352 standard; cDNA; 2235 BP.
XX AC ABK93352;
XX XX 23-AUG-2002 (first entry)
XX XX DNA encoding functionally inactive HOBBIT (HBT) protein #8.
XX KW Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
XX KW pattern formation; plant meristem development; plant yield; drought;
XX KW seedling emergency; root generation; shade avoidance response;
XX KW

```

KW vascular strand formation; patterning; parthenocarpic fruit;
KW cell cycle regulation; endoreduplication; cell division;
KW transgenic plant; plant development; plant morphology; plant physiology;
KW plant biochemistry; gene; ss.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX WO200238599-A2.
XX 16-MAY-2002.
XX 13-NOV-2001; 2001WO-EPL3116.
XX 13-NOV-2000; 2000EP-0870271.
XX 30-NOV-2000; 2000US-250402P.
XX (UYUT-) RIJXSUNIV UTRECHT.
XX Scheres BJG, Bliilou I, Folmer SDH;
XX WPI; 2002-490065/52.
XX Use of plant cdc27B for modulating or mimicking auxin-related effects
PT in plants or plant cells, or for regulation of cell cycle of plant cell
PT
XX
XX Claim 36; Page 170-171; 207pp; English.
XX The invention described the use of a plant cdc27B (also termed HOBBIT
CC (HBT)) useful for modulating or mimicking auxin-related effects in a
CC plant or plant cell. The modulation or mimicking of auxin-related effects
CC results in: altered cell fate and/or altered pattern formation in a plant
CC or plant cell; alteration in the size and/or number of naturally
CC occurring plant meristems; modification of the numbers of organs or
CC tissues, and/or a modification of the rate of organ or tissue emanation
CC from a plant meristem, and/or a modification of the arrangement of organs
CC and/or tissues in a plant; increased plant yield; an increased survival
CC rate of plants; an enhanced survival rate under drought conditions;
CC increase of seedling emergency; mimicking of root generation in tissue
CC cultures; an increased shade avoidance response; altering vascular strand
CC formation and patterning in a plant; or production of parthenocarpic
CC fruits, HBT also regulates the cell cycle of the plant cell results in:
CC modulated endoreduplication in a plant; modulation of sterility in
CC plants; increasing the cell cycle or increase of the rate of cell
CC division; alteration in the size of naturally occurring meristems; a
CC modification in a number of organs or tissues, and/or the modification of
CC the rate of organ or tissue emanation from a plant meristem and/or
CC modification of the arrangement of the organs and/or tissues in a plant
CC and/or an increased plant yield or enhancement of survival rate of
CC plants. A nucleic acid encoding HBT is useful for producing transgenic
CC plants, plant cells or plant tissues and in the regeneration of a plant
CC from a plant cell. HBT nucleic acids and proteins are useful for
CC modifying cell fate, pattern formation, plant development, plant
CC morphology, plant physiology and/or plant biochemistry. This sequence
CC encodes a functionally inactive Arabidopsis thaliana HOBBIT (HBT) or
CC cdc27B protein described in the invention.
XX
XX Sequence 2235 BP; 664 A; 474 C; 484 G; 613 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 0.000149 Length: 2235
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 24 Gaps: 0
US-10-036-492-6 (1-24) x ABK93352 (1-2235)
QY 1 ValAsnLeuClnLeuLeuAlaArgCysTyrLeuSerAsnSerClnAlaTyrSerAlaTyr 20
DB 106 GTTAATTTCAGCTATTAGCCACCAGCTACCTGCAGATAATCAAGCTTACAGTCATAT 165

QY 21 TyrLeuLeuLys 24
DB 166 CATCTGCTAAAG 177
RESULT 11
ABK93354
ID ABK93354 standard; cDNA; 2235 BP.
XX
XX AC ABK93354;
XX
XX DT 23-AUG-2002 (first entry)
XX DE DNA encoding functionally inactive HOBBIT (HBT) protein #10.
XX KW Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
KW pattern formation; plant meristem development; plant yield; drought;
KW seedling emergency; root generation; patterning; shade avoidance response;
KW vascular strand formation; patterning; parthenocarpic fruit;
KW cell cycle regulation; endoreduplication; cell division;
KW transgenic plant; plant development; plant morphology; plant physiology;
KW plant biochemistry; gene; ss.
XX Arabidopsis thaliana.
OS
XX WO200238599-A2.
XX 16-MAY-2002.
XX 13-NOV-2001; 2001WO-EPL3116.
XX 13-NOV-2000; 2000EP-0870271.
XX 30-NOV-2000; 2000US-250402P.
XX (UYUT-) RIJXSUNIV UTRECHT.
XX Scheres BJG, Bliilou I, Folmer SDH;
XX WPI; 2002-490065/52.
XX Use of plant cdc27B for modulating or mimicking auxin-related effects
PT in plants or plant cells, or for regulation of cell cycle of plant cell
PT
XX
XX Claim 36; Page 172-174; 207pp; English.
XX The invention described the use of a plant cdc27B (also termed HOBBIT
CC (HBT)) useful for modulating or mimicking auxin-related effects in a
CC plant or plant cell. The modulation or mimicking of auxin-related effects
CC results in: altered cell fate and/or altered pattern formation in a plant
CC or plant cell; alteration in the size and/or number of naturally
CC occurring plant meristems; modification of the numbers of organs or
CC tissues, and/or a modification of the rate of organ or tissue emanation
CC from a plant meristem, and/or a modification of the arrangement of organs
CC and/or tissues in a plant; increased plant yield; an increased survival
CC rate of plants; an enhanced survival rate under drought conditions;
CC increase of seedling emergency; mimicking of root generation in tissue
CC cultures; an increased shade avoidance response; altering vascular strand
CC formation and patterning in a plant; or production of parthenocarpic
CC fruits, HBT also regulates the cell cycle of the plant cell results in:
CC modulated endoreduplication in a plant; modulation of sterility in
CC plants; increasing the cell cycle or increase of the rate of cell
CC division; alteration in the size of naturally occurring meristems; a
CC modification in a number of organs or tissues, and/or the modification of
CC the rate of organ or tissue emanation from a plant meristem and/or
CC modification of the arrangement of the organs and/or tissues in a plant
CC and/or an increased plant yield or enhancement of survival rate of
CC plants. A nucleic acid encoding HBT is useful for producing transgenic
CC plants, plant cells or plant tissues and in the regeneration of a plant
CC from a plant cell. HBT nucleic acids and proteins are useful for
CC modifying cell fate, pattern formation, plant development, plant
CC morphology, plant physiology and/or plant biochemistry. This sequence
CC encodes a functionally inactive Arabidopsis thaliana HOBBIT (HBT) or
CC cdc27B protein described in the invention.
XX

XX Sequence 2235 BP; 665 A; 475 C; 483 G; 612 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.000149 Length: 2235
 Score: 91.00 Matches: 18
 Percent Similarity: 87.50% Conservativity: 3
 Best Local Similarity: 75.00% Mismatches: 3
 Query Match: 75.21% Indels: 0
 DB: 24 Gaps: 0

US-10-036-492-6 (1-24) x ABK93354 (1-2235)
 QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 Db 106 GTTAATTGGAGCTATTAGCCACCAGCTACCTGCAGATATCAAGCTTACAGTGCATAT 165

QY 21 TyrIleLeuLys 24
 Db 166 CATCTGCTAAAG 177

RESULT 12
 ABK93358
 ID ABK93358 standard; cDNA; 2235 BP.
 AC ABK93358;
 XX
 DT 23-AUG-2002 (first entry)
 DE DNA encoding functionally inactive HOBBIT (HBT) protein #14.
 XX Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
 KW pattern formation; plant meristem development; plant yield; drought;
 KW seedling emergency; root generation; shade avoidance response;
 KW vascular strand formation; patterning; parthenocarpic fruit;
 KW cell cycle regulation; endoreduplication; cell division;
 KW transgenic plant; plant development; plant morphology; plant physiology;
 KW plant biochemistry; gene; ss.
 XX
 OS Arabidopsis thaliana.
 FN WO200238599-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 13-NOV-2001; 2001WO-EP13116.
 XX
 PR 13-NOV-2000; 2000EP-0870271.
 PR 30-NOV-2000; 2000US-250402P.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 XX
 PI Scheres BJG, Blilou I, Folmer SDH;
 XX
 DR WPI; 2002-490065/52.
 XX
 PT Use of plant cdc27B for modulating or mimicking auxin-related effects
 in plants or plant cells, or for regulation of cell cycle of plant cell
 PT
 PS Claim 36; Page 177-178; 207pp; English.
 XX
 CC The invention described the use of a plant cdc27B (also termed HOBBIT
 (HBT)) useful for modulating or mimicking auxin-related effects in a
 CC plant or plant cell. The modulation or mimicking of auxin-related effects
 CC results in: altered cell fate and/or altered pattern formation in a plant
 CC or plant cell; alteration in the size and/or number of naturally
 CC occurring plant meristems; modification of the numbers of organs or
 CC tissues, and/or a modification of the rate of organ or tissue emanation
 CC from a plant meristem, and/or a modification of the arrangement of organs
 CC and/or tissues in a plant; increased plant yield; an increased survival
 CC rate of plants; an enhanced survival rate under drought conditions;
 CC increase of seedling emergency; mimicking of root generation in tissue

CC cultures; an increased shade avoidance response; altering vascular strand
 CC formation and patterning in a plant; or production of parthenocarpic
 CC fruits. HBT also regulates the cell cycle of the plant cell results in:
 CC modulated endoreduplication in a plant; modulation of sterility in
 CC plants; increasing the cell cycle or increase of the rate of cell
 CC division; alteration in the size of naturally occurring meristems; a
 CC modification in a number of organs or tissues, and/or the modification of
 CC the rate of organ or tissue emanation from a plant meristem and/or
 CC modification of the arrangement of the organs and/or tissues in a plant
 CC and/or an increased plant yield or enhancement of survival rate of
 CC plants. A nucleic acid encoding HBT is useful for producing transgenic
 CC plants, plant cells or plant tissues and in the regeneration of a plant
 CC from a plant cell. HBT nucleic acids and proteins are useful for
 CC modifying cell fate, pattern formation, plant development, plant
 CC morphology, plant physiology and/or plant biochemistry. This sequence
 CC encodes a functionally inactive Arabidopsis thaliana HOBBIT (HBT) or
 CC cdc27B protein described in the invention.

SQ Sequence 2235 BP; 664 A; 474 C; 484 G; 613 T; 0 other;

Alignment Scores:
 Pred. No.: 0.000149 Length: 2235
 Score: 91.00 Matches: 18
 Percent Similarity: 87.50% Conservativity: 3
 Best Local Similarity: 75.00% Mismatches: 3
 Query Match: 75.21% Indels: 0
 DB: 24 Gaps: 0

US-10-036-492-6 (1-24) x ABK93358 (1-2235)

QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 Db 106 GTTAATTGGAGCTATTAGCCACCAGCTACCTGCAGATATCAAGCTTACAGTGCATAT 165

QY 21 TyrIleLeuLys 24
 Db 166 CATCTGCTAAAG 177

RESULT 13

ABK93353

ID ABK93353 standard; cDNA; 2313 BP.

XX

AC ABK93353;

XX

DT 23-AUG-2002 (first entry)

XX

DE DNA encoding functionally inactive HOBBIT (HBT) protein #9.

XX

XX Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;

XX pattern formation; plant meristem development; plant yield; drought;

XX seedling emergency; root generation; shade avoidance response;

XX vascular strand formation; patterning; parthenocarpic fruit;

XX cell cycle regulation; endoreduplication; cell division;

XX transgenic plant; plant development; plant morphology; plant physiology;

XX plant biochemistry; gene; ss.

XX

OS Arabidopsis thaliana.

XX

FN WO200238599-A2.

XX

PD 16-MAY-2002.

XX

PF 13-NOV-2001; 2001WO-EP13116.

XX

PR 13-NOV-2000; 2000EP-0870271.

XX

PR 30-NOV-2000; 2000US-250402P.

XX

PA (UYUT-) RIJKSUNIV UTRECHT.

XX

PI Scheres BJG, Blilou I, Folmer SDH;

XX

DR WPI; 2002-490065/52.

XX

PT Use of plant cdc27B for modulating or mimicking auxin-related effects
PT in plants or plant cells, or for regulation of cell cycle of plant cell
PT
XX
XX
PS Claim 36; Page 171-172; 207pp; English.
XX
CC The invention described the use of a plant cdc27B (also termed HOBbit
CC (HBT)) useful for modulating or mimicking auxin-related effects in a
CC plant or plant cell. The modulation or mimicking of auxin-related effects
CC results in: altered cell fate and/or altered pattern formation in a plant
CC or plant cell; alteration in the size and/or number of naturally
CC occurring plant meristems; modification of the numbers of organs or
CC tissues, and/or a modification of the rate of organ or tissue emanation
CC from a plant meristem, and/or a modification of the arrangement of organs
CC and/or tissues in a plant; increased plant yield; an increased survival
CC rate of plants; an enhanced survival rate under drought conditions;
CC increase of seedling emergence; mimicking of root generation in tissue
CC cultures; an increased shade avoidance response; altering vascular strand
CC formation and patterning in a plant; or production of parthenocarpic
CC fruits. HBT also regulates the cell cycle of the plant cell results in:
CC modulated endoreduplication in a plant; modulation of sterility in
CC plants; increasing the cell cycle or increase of the rate of cell
CC division; alteration in the size of naturally occurring meristems; a
CC modification in a number of organs or tissues, and/or the modification of
CC the rate of organ or tissue emanation from a plant meristem and/or
CC modification of the arrangement of the organs and/or tissues in a plant
CC and/or an increased plant yield or enhancement of survival rate of
CC plants. A nucleic acid encoding HBT is useful for producing transgenic
CC plants, plant cells or plant tissues and in the regeneration of a plant
CC from a plant cell. HBT nucleic acids and proteins are useful for
CC modifying cell fate, pattern formation, plant development, plant
CC morphology, plant physiology and/or plant biochemistry. This sequence
CC encodes a functionally inactive Arabidopsis thaliana HOBbit (HBT) or
CC cdc27B protein described in the invention.
XX
SQ Sequence 2313 BP; 682 A; 493 C; 494 G; 644 T; 0 other;

Alignment Scores:
Pred. No.: 0.000155 Length: 2313
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 24 Gaps: 0

US-10-036-492-6 (1-24) x ABK93353 (1-2313)

QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 106 GTTAATTGCGCTATTAGCCACCTACCTGCAGATAATCAAGCTTACAGTGCATAT 165

QY 21 TyrIleLeuIys 24
Db 166 CAPCTGCTAAG 177

RESULT 14
AAC50257
ID AAC50257 standard; DNA; 2316 BP.
AC AAC50257;
XX
XX 19-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 64158.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX

PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126364.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128334.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132883.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.

```
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142809.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143543.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144322.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 0.000155 Length: 2316
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 21 Gaps: 0

US-10-036-492-6 (1-24) x AAC50257 (1-2316)
Qy 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 106 GTTAATTGCGAGCTATTAGCCACAGCTACCTGCAGATAATCAAGCTTACAGTGCATAT 165
Qy 21 TyrIleLeuLys 24
Db 166 CATCTGCTAAAG 177

RESULT 15
ABK93341
ID ABK93341 standard; cDNA; 2480 BP.
XX AC ABK93341;
XX XX
XX 23-AUG-2002 (first entry)
XX DE
XX DNA encoding cell cycle regulation protein HOBBIT (HBT) #1.
KW Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
KW pattern formation; plant meristem development; plant yield; drought;
KW seedling emergency; root generation; shade avoidance response;
KW vascular strand formation; patterning; parthenocarpic fruit;
KW cell cycle regulation; endoreduplication; cell division;
KW transgenic plant; plant development; plant morphology; plant physiology;
```

KW plant biochemistry; gene; ss.

XX Arabidopsis thaliana.

XX WO200238599-A2.

XX 16-MAY-2002.

XX 13-NOV-2001; 2001WO-EP13116.

XX 13-NOV-2000; 2000EP-0870271.

XX 30-NOV-2000; 2000US-250402P.

XX (UUT-) RIJXSUNIV UTRECHT.

XX Scheres BJG, Blilou I, Folmer SDH;

XX WPI; 2002-490065/52.

XX Use of plant cdc27B for modulating or mimicking auxin-related effects
PT in plants or plant cells, or for regulation of cell cycle of plant cell

PT

XX Claim 31; Page 140-142; 207pp; English.

XX The invention described the use of a plant cdc27B (also termed HOBbit
XX (HBT)) useful for modulating or mimicking auxin-related effects in a
CC plant or plant cell. The modulation or mimicking of auxin-related effects
CC results in: altered cell fate and/or altered pattern formation in a plant
CC or plant cell; alteration in the size and/or number of naturally
CC occurring plant meristems; modification of the numbers of organs or
CC tissues, and/or a modification of the rate of organ or tissue emanation
CC from a plant meristem, and/or a modification of the arrangement of organs
CC and/or tissues in a plant; increased plant yield; an increased survival
CC rate of plants; an enhanced survival rate under drought conditions;
CC increase of seedling emergency; mimicking of root generation in tissue
CC cultures; an increased shade avoidance response; altering vascular strand
CC formation and patterning in a plant; or production of parthenocarpic
CC fruits. HBT also regulates the cell cycle of the plant cell results in:
CC modulated endoreduplication in a plant; modulation of sterility in
CC plants; increasing the cell cycle or increase of the rate of cell
CC division; alteration in the size of naturally occurring meristems; a
CC modification in a number of organs or tissues, and/or the modification of
CC the rate of organ or tissue emanation from a plant meristem and/or
CC modification of the arrangement of the organs and/or tissues in a plant
CC and/or an increased plant yield or enhancement of survival rate of
CC plants. A nucleic acid encoding HBT is useful for producing transgenic
CC plants, plant cells or plant tissues and in the regeneration of a plant
CC from a plant cell. HBT nucleic acids and proteins are useful for
CC modifying cell rate, pattern formation, plant development, plant
CC morphology, plant physiology and/or plant biochemistry. This sequence
CC encodes an Arabidopsis thaliana HOBbit (HBT) or cdc27B protein described
CC in the invention.

XX Sequence 2480 BP; 737 A; 527 C; 533 G; 683 T; 0 other;

Alignment Scores:

Pred. No.:	0.000169	Length:	2480
Score:	91.00	Matches:	18
Percent Similarity:	87.50%	Conservative:	3
Best Local Similarity:	75.00%	Mismatches:	3
Query Match:	75.21%	Indels:	0
DB:	24	Gaps:	0

US-10-036-492-6 (1-24) x ABK93341 (1-2480)

Qy	1	ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr	20
Db	121	GTATATTGGAGCTATTAGCCACGAGCTACCTGCAGATATATCAAGCTTACAGTGCATAT	180
Qy	21	TyrIleLeuLys	24
Db	181	CATCTGCTAAAG	192

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: February 2, 2004, 13:20:50 ; Search time 6282 Seconds
(without alignments)
156.293 Million cell updates/sec

Title: US-10-036-492-6
Perfect score: 121
Sequence: 1 VNLQLLARCYLSNQAYSAYILK 24

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model -DRV=xlh
-Q=/cgn2_1/USPTO_spool/US10036492/runat_02022004_085033_13457/app.query.fasta_1.199
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HRAFSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10036492 @CGN 1.1 3508 @runat_02022004_085033_13457 -NCFU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPCP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.ly.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	100.0	2401	6	AX068991 Sequence
2	121	100.0	2434	6	AX068986 Sequence
3	121	100.0	75289	8	AB023046 Arabidops
C 4	121	100.0	91714	8	ATAC001645
C 5	121	100.0	135457	2	AC084820 Arabidops
6	91	75.2	486	6	AX462180 Sequence
7	91	75.2	498	6	AX462181 Sequence
8	91	75.2	1557	6	AX462198 Sequence
9	91	75.2	1915	8	AY128780 Arabidops
10	91	75.2	2115	6	AX462197 Sequence
11	91	75.2	2169	6	AX462196 Sequence
12	91	75.2	2220	6	AX068992 Sequence
13	91	75.2	2235	6	AX462179 Sequence
14	91	75.2	2235	6	AX462193 Sequence
15	91	75.2	2235	6	AX462195 Sequence
16	91	75.2	2235	6	AX462199 Sequence
17	91	75.2	2313	6	AX462194 Sequence
18	91	75.2	2480	6	AX462178 Sequence
19	91	75.2	2672	8	AY062470 Arabidops
20	91	75.2	2877	8	AY487669 Arabidops
21	91	75.2	4577	6	AX462177 Sequence
22	91	75.2	4577	6	AX462186 Sequence
23	91	75.2	4577	6	AX462187 Sequence
24	91	75.2	4577	6	AX462188 Sequence
25	91	75.2	4577	6	AX462189 Sequence
26	91	75.2	4577	6	AX462190 Sequence
27	91	75.2	4577	6	AX462191 Sequence
28	91	75.2	4577	6	AX462192 Sequence
29	91	75.2	97814	8	AC006081 Arabidops
30	87	71.9	140327	2	AF003938 Oryza sat
31	87	71.9	173297	2	AF003539 Oryza sat
32	63	52.1	2995	5	BC049466 Dario rer
C 33	61	50.4	73450	2	AC137580 Homo sapi
C 34	61	50.4	149041	9	AC044782 Homo sapi
C 35	61	50.4	179789	9	AC069543 Homo sapi
C 36	61	50.4	180915	8	GFADJ10552
37	59	48.8	314	6	BD025555 Sequence
38	59	48.8	1864	10	BC023187
39	59	48.8	2472	6	I91768 Sequence 2
40	59	48.8	2592	9	HSCDC27
41	59	48.8	3309	9	BC011656 Homo sapi
42	59	48.8	3320	9	BC011656 Homo sapi
C 43	59	48.8	53128	2	AC068237 Homo sapi
44	59	48.8	67150	2	AC025606 Homo sapi
45	59	48.8	102064	9	AC002558 Homo sapi

ALIGNMENTS

RESULT 1

```

AX068991
LOCUS AX068991 2401 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 14 from Patent WO0102430.
ACCESSION AX068991
VERSION AX068991.1 GI:12578821
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Hemerly,A.S., Ferreira,P.C. and Rombauts,S.
TITLE Arabidopsis thaliana cdc7 and cdc27 homologs
JOURNAL Patent: WO 0102430-A 14 11-JAN-2001;
CropDesign N.V. (BE) ; Universidade Federal do Rio de Janeiro (BR)
FEATURES
source
1..2401
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 747 a 466 c 514 g 674 t
ORIGIN
Alignment Scores:
Pred. No.: 2,26e-10 Length: 2401
Score: 121.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-036-492-6 (1-24) x AX068991 (1-2401)
Qy 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 109 GTGAACCTGCAATTGTAGCCAGGTGTTACTTGATGACAGTCAGCTTATAGTCATAT 168
Qy 21 Tyrlleululs 24
Db 169 TATATCCTTAAA 180

RESULT 2
LOCUS AX068986 2434 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 9 from Patent WO0102430.
ACCESSION AX068986
VERSION AX068986.1 GI:12578820
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Hemerly,A.S., Ferreira,P.C. and Rombauts,S.
TITLE Arabidopsis thaliana cdc7 and cdc27 homologs
JOURNAL Patent: WO 0102430-A 9 11-JAN-2001;
CropDesign N.V. (BE) ; Universidade Federal do Rio de Janeiro (BR)
FEATURES
source
1..2434
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 755 a 475 c 520 g 684 t
ORIGIN
Alignment Scores:
Pred. No.: 2,29e-10 Length: 2434
Score: 121.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

```

```

DB: 6 Gaps: 0
US-10-036-492-6 (1-24) x AX068986 (1-2434)
Qy 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 109 GTGAACCTGCAATTGTAGCCAGGTGTTACTTGATGACAGTCAGCTTATAGTCATAT 168
Qy 21 Tyrlleululs 24
Db 169 TATATCCTTAAA 180

RESULT 3
LOCUS AB023046 75289 bp DNA linear PLN 27-DEC-2000
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: MYA6.
ACCESSION AB023046 BAO00014
VERSION AB023046.1 GI:4220645
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (sites)
Sato,S., Nakamura,Y., Kaneko,T., Katoh,T., Asamizu,E. and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 3. I.
Sequence features of the regions of 4,504,864 bp covered by sixty
P1 and TAC clones
DNA Res. 7 (2), 131-135 (2000)
MEDLINE 20277480
PUBMED 10819329
REFERENCE
2 (bases 1 to 75289)
Sato,S., Nakamura,Y., Kaneko,T., Katoh,T., Asamizu,E. and Tabata,S.
Direct Submission
Submitted (01-FEB-1999) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1332-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/sgd/graph.cgi?c=MYA6
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3//),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremli.zool.kiastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE//).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is MSU1 and the 3' clone is MDC8.
FEATURES
source
1..75289
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MYA6"
/clone_lib="Mitsui P1"
complement(join(4483..4904,4991..5106,5192..5636,
6082..6469))
/note="gene_id:MYA6.1
unknown protein"
/codon_start=1
CDS

```

```

/evidence=not experimental
/protein_id="BAB01260.1"
/db_xref="GI:9279703"
/translation="MERGMNGQSLWHSQSPKPTTMDRLASRRRPHSDADLASGE
SGDESCKPHYLILASNLGRHGWPCLLIALFLVLLFLISVAHSHSFVCIS
RFDPAARIGFDLGLSDFGALGQWCRSGKGEVMTSKOLLGLEEFVPIYETRP
IKNNYGMGPDHSGFLMFAWPKDPMIESGAPKGHSTWLRQAMPDTPMI SLTPRH
PEKYLKRGADYDNGCYFYAGKQWDFDGSVDKWNLRKHGCTDLSRVIVFDDHNSL
KRUKQALACAGFRHLIPDNDYDTGTGDHYSLRQICDOSHIRGGGHSFCFQSDAARMSEK
RKPFWEAVATEELCGRGTWGVKGEMRDDFNHTPTPSYNGHFQNSRYVESILDVY
WEUPPVPAGSPSLTHOSRVPDAPRPIIVADGKHRLFORIGLRDCKSVFNQYTMWYLE
ISKPGS"
complement(7397..8479)
/notes="gb|AAF25964.1
gene_id:MYA6.2
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01261.1"
/db_xref="GI:9279704"
/translation="MSKFLPEELAEILVRLSKMDLARFCVCKTWRDLINDPGFTET
YRDMSPAKFSFYDKNFMVLDVEKGHPVITNKLDPLDQSDIDESTCVLHCDGTLCTV
LKHTLWNPFSKQKIVPNPGIYQDSNLTGFGYDQVHDYKVFIFDLDDVSTAHV
PEPTKSGWBSLQISYDPMHYRBRGTFELDYLWYAYRSADRFILCNLSYHEK
LPLVYNGQTSWGLGVTQKLCITTEYEMCKEIRISVMEKTSWAKIISLSWFSIS
VQDRIYDQVEFVSFRKNDLVVTFGTYNDFHEPEPRTKKOMFLYKYNKENSEVR
FCNPLAQLRFLCEVETLKNRFI"
join(10577..10657,10791..10941,11019..11131,11203..11292,
11366..11429,11555..11619,11726..11932)
/notes="gene_id:MYA6.3
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01262.1"
/db_xref="GI:9279705"
/translation="MENLKHWRCSSSHOTPLSFKNQKQIGOGOREVTFHVSPLP
AIYPLDKTEAFQSVLGNNDKPLQFSTLAEMGTEKISFVTPKTEHLTVMLKLE
NNESVQAQNILOVLSPEKQSLFSNVROALKNRPVFLRGLGCKSGSLDKTVLYA
PVEBVGHVDAENVGAFGAKSLKHLATLNNASYRDXKSKMDTFDAREIHKEF
ENKDWGTLIREAHISORYKIDNGYFHCASLPFPHK"
join(13499..13671,13836..13881,14227..14264,14340..14419,
14566..14623,14732..15011,15103..15215,15295..15360,
15467..15572,15667..15731,15855..15895,15984..16167)
/notes="gene_id:MYA6.4
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01263.1"
/db_xref="GI:9279706"
/translation="MIFQSNKTONFIFLAMVLEKSFDFVLFLFMHVHVCVGLIT
VYGLRVDFVNRNRSKSDMDGYKKQKMGASLIKPIRKEGTQWKLBEEMGVKII
LPSENKDHISIEGSDVCTKASKRIATIIDVWSPSLDYSHFVSLPIAHPELVLD
KLNVQNSILGHSIADKQDDQANRATTSVAVDLKANSETNQNVGKISIPVSPPP
KAKSSTLDLGLERKIFIKPSTFHTLVVWLKNKORVNAACDVLSKIFPSPVMDAL
DNKPVFTLRGLGDCMRGFLDKTRVLYAPVEIGDEGLRLACRILFDITDAFVKAGL
VLEKDAQSKLHYVTVNVAHRKRNKKKMTFEDARETHKFGNEDMGXYLQBAH
LSORFVFDQNGYVRCGSIIPFGQRA"
join(18398..18524,18628..18878,19278..19652)
/notes="gb|AAC49281.1
gene_id:MYA6.5"
/codon_start=1
/evidence=not experimental
/product="delta tonoplast intrinsic protein"
/protein_id="BAB01264.1"
/db_xref="GI:9279707"
/translation="WAGVARGSPDPSFLASRLAYLAEEFTILLFVAGVGSIAVAK
LTSDAALDTGLVAIVCHGFALPVAIGANTSGHVPNAVTFGLVAGQITVIGV
FYWAQLIGSTAAFLKRYVTGLAVFTHSVAAGISGIEGVNEIITFALVTVIAT
AADPKKSLGTAPLATGLIVGANILAAFPFGSGMNPARSFGFAVAAAGDFSGHWYV
VGPIIGGGLAGLTYGVNFMGSSSEHVPLASADF"
complement(join(19955..20047,20150..20212,20303..20365,
20457..20513,20611..20721,20957..21184))
/notes="gene_id:MYA6.6

```

```

unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01265.1"
/db_xref="GI:9279708"
/translation="WGSVOLSGSLVGLASLPNHFSFKTKLKNKNSYFFRKHNAART
KTSRAISTAPASQPAADPEPFAVDPAFVHSLVLLPDGTFDWHRRNAGQOKURL
LDSNIELYGPYKSLNCGAGVGTATCVMVEIVNGKELNLPRTDIEKKELKRPKWRL
ACQTVNGPNDSTGLVVIQOLPEWKAHEMNIPKPNDDDLST"
join(22261..22656,22735..22824,23038..23092,23187..23230,
23347..23742,23833..24098,24182..24326,24415..24465,
24561..25553,25635..25982)
/notes="gb|AAF29402.1
gene_id:MYA6.7
strong similarity to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01266.1"
/db_xref="GI:9279709"
/translation="MLTSSMPQNLSPGFSPLKSSPFLIIRPSPLYPIIPASSPAP
SRPRPTAGYRRSGSPPRKRWSSFEQKRGSPMEKDKAISFNHSDSDSEFNKRRR
EGLDKVDPKKNLKTNTLTNTIAYQILGTMDTQDTSPSVLLFDQKRFIFNA
GEVDHIFLSRVCSETAGLPGLLTLAGIGEGSLSVNWPGLSDLYLVAMRSPTPA
AMWTRSPGSLNIDSAPOLGSKPKDDAVLVDDVVKISAILBPSRLBESGSKP
GETAVIYCELPKELKGFEDPKKAMALGIRAGKYSLYLGOSGOSDKDKDITVHPSDVM
GSPVCPVLLVDQFTESHAEELLISPMKTYISCLDSTGAKLVNGLIILHSASVT
NSSTYRVMKPFSAQHILAGHEAKNMEFFILRASSRTARLNLCPQFFPAGFWSH
QDNHNSINPTSLSKCFDNLGSEISAENLKEFTLRPHGLNVDGRSSISRLTAURVMD
ELLSEIPISSTKEEIKOLMNGQHNMMIEEPWLGESTVPSCLNIRDDMEI VLLGT
SSQSPKYRNVTAIYIDFSGSILLDCGEGTLQKKRYGLEGADEAVRLNLCIWS
HIHADHTGLARILARRRELLGLAHEPAIVVGRPSLKNFLCAVORLEDLNVEFLDCR
NLTWTSWASVETSPKNTSSNAGSLFSKSLMQSIYKRPSPSLTNSALPLKK
LXVLGEMGLEHLISFPVVPQCPQAFVSLKAAERKNIAAGDEIPGKMYVSGDTRCPK
MYEASGLVLIHEATFEDALVEEAVAKNHTTKEAIRKVGSSAGYRIVRLIHFQSRYP
KIPVDESHMNTCIAFDMSINWADHLVLPKILFYFKTLFRNQVVEEETDDDD
SLIRKQVSPFTIN"
join(26565..26771,27055..27792,27965..28120,28211..29212,
29363..29404)
/notes="gb|AAB63649.1
gene_id:MYA6.8
strong similarity to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB01267.1"
/db_xref="GI:9279710"
/translation="MDTSRAVESYWRSMIDAVTSDKQVAPVYKLEICDLRSSH
VSVIRFSEPIKLNDKSPVQKALRLIKYANGKSGSEFRERQORSVAVRNLFHY
KHPDPLKGDALNAKAVRTHETTSIAIFSENGTKPAAPESINRRIGFGNTNFOVPS
NDNKSFLSEVVGISIKQISNFAQGHLPKQNGNSYRGPNLHSLTMENFNS
RYDPVKLGDNNGYTSKNTGSGWGHASGEASASVRSKTEEKLEETITVSSG
RYQPTDALHVFILAAKMDVALSIALDGKLSHPMWQVMKALCVLEAILRKDEE
NFSI VHTYFSENDAIORCAESPOSSLEKANKVLSLLNGQSSGLSSSDNTVKREA
AVDPLDIDTSGSDTLNNAIITGTVATAGPLMDDMDFGSDSDIGLSSSEKTD
DPFADVSPHNEEKESADDFSGMTVGEKSAVGNHVDPDLDFMGSGTAKLEAEKDA
KNINDMGFSFIDENNSKQSSSTLPQDLFAMPSTTSHQAPENFVGILIGSQNPGF
TQNTLPGVVPFPFQGMNPAFASQPLNAYAMASLLAQOQYLGNSFQOQGNL
NAQGNGLSVMTSGNQSLPDIPOPNFGQAFTSTMNGSKKEDTRAPDFISVSFAI
LILIFDSKVSICHLSDYMDHLTSARDTKRS"
join(31165..31929)
/notes="gene_id:MYA6.9"
/codon_start=1
/evidence=not experimental
/product="transcription factor TINY-like protein"
/protein_id="BAB01268.1"
/db_xref="GI:9279711"
/translation="MSHTTKRSKYQTRVOKWTSLSNASSAGTSSSSQSDATTTTST
HLSEEAEPNNNTEKKRRDSSASSSMOHPPYVGRVMSRWGKWSEIRQPKKTR
TWLGTFTVADVAARADHVAALTIKSSSAVLNFPPELASLFRPASSSHIDIQTAAAEA

```

Alignment Scores:
Pred. No.: 8,24e-09 Length: 75289
Score: 121.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 8 Gaps: 0

US-10-036-492-6 (1-24) x AB023046 (1-75289)

QY 1 ValAsnLeuGlnLeuLeuAlaAGCyStyLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 |||||
 Db 42990 GTGAACCTGCATGTTAGCCAGTGTACTTACTGAGTAACAGTCAGCTTATAGTCAATAT 43049
 |||||

QY 21 TyrIleLeuIlys 24
 |||||
 Db 43050 TATATCCTTAAA 43061
 |||||

RESULT 4
 ATAC001645/C 91714 bp DNA linear PLN 30-OCT-2002
 LOCUS Arabidopsis thaliana chromosome III BAC T02004 genomic sequence,
 DEFINITION complete sequence.
 AC001645
 AC001645.1 GI:2062153
 KEYWORDS HTG.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 91714)
 Rounsley, S.D., Lin, X., Kechum, K.A., Phillips, C.A., Brandon, R.C.,
 Fuhrmann, J.L., White, O., Kerlavage, A.R., Adams, M.D.,
 Somerville, C.R. and Venter, J.C.
 Arabidopsis thaliana chromosome III BAC T02004 genomic sequence
 Unpublished
 2 (bases 1 to 91714)
 Rounsley, S.D.
 DIRECT SUBMISSION
 Submitted (22-APR-1997) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 3 (bases 1 to 91714)
 Rounsley, S.D.
 DIRECT SUBMISSION
 Submitted (22-JUL-1997) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 Address all correspondence to:
 Steve Rounsley
 The Institute for Genomic Research
 9712 Medical Center Dr.,
 Rockville, MD 20850,
 USA
 e-mail: rounsley@tigr.org

BAC clone T02004 is from Arabidopsis chromosome III and mapped
 close to the molecular marker m28.
 The orientation of the sequence is from SP6 to T7 end of the BAC
 clone.

Genes were identified by a combination of three methods: Gene
 prediction programs including GRAIL (available by anonymous ftp
 from arthur.epm.ornl.gov), Genefinder (Phil Green, University of
 Washington), Genscan (Chris Burge,
<http://www.cbs.dtu.dk/netpgene/chanetpgene.html>), and NetPlantGene
<http://www.cbs.dtu.dk/netpgene/chanetpgene.html>). Searches of the
 complete sequence against a peptide database and the Arabidopsis
 EST database at TIGR (<http://www.tigr.org/tdb/at.html>).
 Annotated genes are named to indicate the level of evidence for
 their annotation. Genes with similarity to other proteins are named
 as isoforms. Genes without significant peptide similarity but with
 EST similarity are named as 'unknown' proteins. Genes without
 protein or EST similarity, that are predicted by more than two gene
 prediction programs over most of their length are annotated as
 'hypothetical' proteins. Genes encoding tRNAs are predicted by
 tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-se/>).
 Simple repeats are identified by RepeatMasker (Arian Smit,
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Regions of
 genomic sequence that are not annotated as genes but have predicted

exons by GRAIL are annotated as misc features.

FEATURES
 source
 Location/Qualifiers
 1..91714
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="III"
 /map="m28"
 /clone="T02004"
 complement(100..125)
 /note="exon predicted by xgrail, quality
 marginal_shadowexon"
 <228..851
 /gene="T02004.1"
 /note="gene model predicted by genscan, 5' partial"
 <228..851
 /gene="T02004.1"
 /note="T02004.1"
 /note="hypothetical protein"
 /codon_start=1
 /protein_id="AB063628.1"
 /db_xref="GI:2062154"
 /translation="ERLDDSERSEIHSKRIISVEKOSNNNNAYDTSPIKVIDTVK
 TKSRKMNVAVSECDDFIYQAKDFWSPGKCKPFTACNTPRSSVANNYYIT
 PPSFASVCDACFRSYPGLMTPFSKAKVRSHSAFQRPDRKRLSLDEIM
 AASSVSGVMVQPOQPOQTQOQKRSQSDYHQPRQNETDPRFYN"
 complement(1293..1358)
 /note="exon predicted by xgrail, quality
 excellent_shadowexon"
 complement(1297..4907)
 /gene="T02004.2"
 complement(join(2097..2143,2248..2360,2454..2503,
 2589..2659,2749..2818,2908..3005,3127..3193,3267..3344,
 3434..3601,3713..3880,3967..4263,4346..4579,4869..4907))
 /gene="T02004.2"
 complement(join(2097..2143,2248..2360,2454..2503,
 2589..2659,2749..2818,2908..3005,3127..3193,3267..3344,
 3434..3601,3713..3880,3967..4263,4346..4579,4869..4907))
 /gene="T02004.2"
 /codon_start=1
 /product="mitochondrial processing peptidase alpha subunit
 precursor isoform"
 /protein_id="AA063629.1"
 /db_xref="GI:2062155"
 /translation="MRTAASRAKALKGILNHNFRASRYASSAVATSSSSSSWLSGG
 YSSLPSPMNIPLAGVSLPPLSDHVEFSPKLTTLNGLTIATENSPNRAISGLIYD
 CGSIYTFPQFGATHLLERNAFKSTLNRSFRLVREIEAGTNSASASRQMGYITD
 AKTYPEMVVEVLIDSVNPAPFLDWEVNEELRKVEIGEFATNPMGLLEAVHSAGY
 SGALANPLYAPSAITGLTGEVLENFVFNITASRMVLAASGVDBELLKXVEPLISG
 LPNVPAPBPSQVVGDPDPCHTGHEATHPALAFVPGWNNKEAIIATVLOMLGG
 GSPSAGPGKGMHSLYLRLLNOHQOQSCATFVNNNTGLFGIGCTSPFASQGI
 ELVASENNVADGKVNQKHLDRAKATKSAIKNLSRMTAAEDIGQILTYGERKPV
 DQFLTKVDQLKDIADFTSKYTKTLTATFGDVLNVPISYDSVSRFR"
 complement(6086..6130)
 /note="exon predicted by xgrail, quality
 marginal_shadowexon"
 6398..6436
 /rpt_family="POLY A"
 complement(6535..6567)
 /rpt_family="AT-rich"
 complement(6782..6887)
 /note="exon predicted by xgrail, quality
 marginal_shadowexon"
 7046..8876
 /gene="T02004.3"
 join(7046..7298,7382..7843,7929..8378,8469..8876)
 /gene="T02004.3"
 join(7104..7298,7382..7843,7929..8378,8469..8717)
 /gene="T02004.3"
 /codon_start=1
 /product="jasmonate inducible protein isolog"

misc_feature
 repeat_region
 repeat_region
 misc_feature
 gene
 mRNA
 CDS


```
/protein_id="AAB63630.1"
/db_xref="GI:2062156"
/translation="MAKKLEAAGRGSEEDDGGAVENKVVVGQDSGVVVVVFYD
EKQKLVSHHGKQTLTGTEEFVDPEDYITSVKLYEKLFGSPLEIVTALLFKTFKG
KTSQPLGLTSGEARELGKGVGFHGSDDLHSVGVIIIPSTPLTPVPSGLIKLE
AAGRGDWDVDDGVAVNVKVVVGQDSGVVVKFDYDKDKIVLSLEHGKQTLIGTE
EPEIDEDYITVYKLYEKLFGSPLEIVTALLFKTFKGTSQPLGLTSGEARELGK
IVGHFTSSILHSLGAVIIPSTPLTPSNTIPAGQSGVAVDDGVHSDVKKIYVG
QGSCTVYFKADYKAKVPKLGSDHGKTLGAEFVLGPDVYVAVSGYDKIIFSDV
APAVLSLKFNKTRKTSIPYGLEGTFFVLEKDKHKIVGVFGAGELYKLGVNAPIA
K"
complement(9017..9070)
/notes="exon predicted by xgrail, quality
marginal shadowexon"
complement(9242..11792)
/gene="T02004.4"
complement(join(9242..9539,9634..10095,10178..10630,
10728..11534,11638..11792))
/gene="T02004.4"
complement(join(9291..9539,9634..10095,10178..10630,
10728..11534,11638..11792))
/gene="T02004.4"
/codon_start=1
/product="jasmonate inducible protein isolog"
/protein_id="AAB63631.1"
/db_xref="GI:2062157"
/translation="MSWDDGSHAKVKKVQLTFDEIIYSIQVYDGTALQSLRGSVG
PKSAFTLAPDEIVTALSAYGSLSTOEIVTALTFTNTKSYGPKTSKGFQISAPEA
TGKISGLFTSGNVLNIDVHSPIPTGTTGGTGTGTGTGTGTGTGTGTGTGTGTGT
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
TKIVASGGGSIQVAFDYVKGVTQGVHLHGKQKQKQKQKQKQKQKQKQKQKQKQ
YETVMLGIQKTNLYEYVSIYFEPSTDTKFLQVQDKKILGFHGFAGNHNYSIGAY
FVPSKSTVPSTPLKLTARGCGTGAVDGSDHDDKVKYVVGQDGVAAVFEYKNG
SQVFGDERTRTLLGFEPELESDEIITSYGGYKRNKFGVDVTTLLFKKSKNTIA
GPFIVSGTFFKFGKGVKITGFHGRAGEYNAIGALPAGTPTPLTPATQSKLEGA
GSGAGTLWDGADGVKVGSGQAGQDGVSFVYKAGQVGEKGHGTLLGLFEF
ELDYPSTYITAVDGTVDIAIENEPINVMFLRTNKRVSIPFGIGAGTAFEKDGOKI
VGFHGRAGDILLHFGVHVAFITK"
11880..11992
/notes="exon predicted by xgrail, quality
excellent shadowexon"
12050..12135
/rpt_family="TAGA)n"
12541..12563
/rpt_family="AT_rich"
13248..13301
/rpt_family="AT_rich"
complement(13383..13458)
/rpt_family="AT_rich"
complement(13499..13623)
/rpt_family="(TAA)n"
complement(13589..13701)
/rpt_family="(TAA)n"
complement(13931..14021)
/rpt_family="AT_rich"
complement(14268..14288)
/notes="exon predicted by xgrail, quality good"
complement(15021..16259)
/gene="T02004.5"
complement(join(15021..15272,15386..15841,16026..16259))
/gene="T02004.5"
complement(join(15021..15272,15386..15841,16026..16220))
/gene="T02004.5"
/codon_start=1
/product="jasmonate inducible protein isolog"
/protein_id="AAB63632.1"
/db_xref="GI:2062158"
/translation="MAKQVZAGGAGGASWDDGVHGVKVVVGQDGVSSINVVYA
KSDQVGEHGHKTLGTFETFEVADDDIYAVQVYDYNVFGQDSIIITSIFNFKG
```

Alignment Scores:
Pred. No.: 1.01e-08 Length: 91714

```
Score: 121.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-036-492-6 (1-24) x ATAC001645 (1-91714)

QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrIeuSerAenSerGlnAlaTyrSerAlaTyr 20
Db 74667 GTGAACCTGCAATGTTAGCCAGGTGTACTTGAGTAACAGCTTATAGTGCATAT 74628

QY 21 TyrlleleuLys 24
Db 74627 TATATCCTTAAA 74616

RESULT 5
AC084820/c
LOCUS AC084820 135457 bp DNA linear HTG 08-MAR-2003
DEFINITION Arabidopsis thaliana chromosome 1 clone T2P3 strain Columbia, ***
SEQUENCING IN PROGRESS ***, 33 unordered pieces.
AC084820
VERSION AC084820.13 GI:28867009
KEYWORDS HTG; HTGS PHASE1.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 135457)
AUTHORS Town,C.D., Haas,B., Tallon,L.J., Rooney,T., Ciecko,A.,
Utterback,T., Vanaken,S., Feldblyum,T., White,O. and Fraser,C.M.
Arabidopsis thaliana BAC T2P3
Unpublished
2 (bases 1 to 135457)
AUTHORS Town,C.D. and Kaul,S.
Direct Submission
TITLE Submitted (21-NOV-2000) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA, cdtown@igr.org
3 (bases 1 to 135457)
AUTHORS Town,C.D., White,O. and Fraser,C.M.
JOURNAL Direct Submission
TITLE Submitted (08-MAR-2003) The Institute for Genomic Research, 9712
COMMENT On Mar 6, 2003 this sequence version replaced gi:28626637.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 32150: contig of 32150 bp in length
* 32151 32250: gap of unknown length
* 32251 37229: contig of 5079 bp in length
* 37229 37429: gap of unknown length
* 37430 42503: contig of 5074 bp in length
* 42503 42603: gap of unknown length
* 42603 44061: contig of 1458 bp in length
* 44061 44161: gap of unknown length
* 44161 48220: contig of 4059 bp in length
* 48220 48320: gap of unknown length
* 48320 53694: contig of 5374 bp in length
* 53694 53794: gap of unknown length
* 53794 54165: contig of 371 bp in length
* 54165 54265: gap of unknown length
* 54265 56819: contig of 2554 bp in length
* 56819 56920: gap of unknown length
* 56920 57111: contig of 792 bp in length
* 57111 57811: gap of unknown length
* 57811 59058: contig of 1247 bp in length
* 59058 59158: gap of unknown length
```

* 59159 60027: contig of 869 bp in length
* 60028 60127: gap of unknown length
* 60128 60876: contig of 749 bp in length
* 60877 60976: gap of unknown length
* 60977 61222: contig of 246 bp in length
* 61223 61902: contig of 580 bp in length
* 61903 62002: gap of unknown length
* 62003 62977: contig of 975 bp in length
* 62978 63077: gap of unknown length
* 63078 63886: contig of 809 bp in length
* 63887 63986: gap of unknown length
* 63988 64996: contig of 710 bp in length
* 64997 64796: gap of unknown length
* 64797 65406: contig of 610 bp in length
* 65407 65507: gap of unknown length
* 65508 66479: contig of 973 bp in length
* 66480 66579: gap of unknown length
* 66580 68666: contig of 287 bp in length
* 68667 68966: gap of unknown length
* 68967 68583: contig of 1617 bp in length
* 68584 68683: gap of unknown length
* 68684 69531: contig of 848 bp in length
* 69532 69631: gap of unknown length
* 69632 69808: contig of 177 bp in length
* 69809 69908: gap of unknown length
* 69909 70062: contig of 154 bp in length
* 70063 70162: gap of unknown length
* 70163 76498: contig of 6336 bp in length
* 76499 76598: gap of unknown length
* 76599 79392: contig of 2794 bp in length
* 79393 79492: gap of unknown length
* 79493 85493: contig of 6001 bp in length
* 85494 85593: gap of unknown length
* 85594 87441: contig of 1848 bp in length
* 87442 92881: contig of 5340 bp in length
* 92882 92981: gap of unknown length
* 92982 93835: contig of 854 bp in length
* 93836 93935: gap of unknown length
* 93936 94804: contig of 869 bp in length
* 94805 94904: gap of unknown length
* 94905 96190: contig of 1286 bp in length
* 96191 96290: gap of unknown length
* 96291 135457: contig of 39167 bp in length.

FEATURES

Location/Qualifiers
source
1..135457
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="T2P3"

BASE COUNT 42738 a 22476 c 23879 g 43123 t 3241 others
ORIGIN

Alignment Scores:

Pred. No.: 1.52e-08 Length: 135457
Score: 121.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-036-492-6 (1-24) x AC094820 (1-135457)

Qy 1 ValAsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 70569 GTGACCTGCAATGTTAGCCAGGTTACTTGTAGTACACGCTTATAGTCATAT 70510

Qy 21 TyrIleLeuLys 24

Db 70509 TATATCCCTAAA 70498

RESULT 6

AX462180
LOCUS AX462180 486 bp DNA linear PAT 09-JUL-2002
DEFINITION Sequence 4 from Patent WO0238599.
ACCESSION AX462180
VERSION AX462180.1 GI:21727743
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1 Scheres,B.J., Bllou,I. and Folmer,S.D.
AUTHORS A plant development regulating gene and its uses
TITLE Patent: WO 0238599-A 4 16-MAY-2002;
JOURNAL Universiteit Utrecht (NL)

FEATURES

Location/Qualifiers
source
1..486
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

BASE COUNT 132 a 108 c 102 g 144 t
ORIGIN

Alignment Scores:

Pred. No.: 4.71e-06 Length: 486
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 6 Gaps: 0

US-10-036-492-6 (1-24) x AX462180 (1-486)

Qy 1 ValAsnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 106 GTTAATTTGCAGCTATTAGCCAGGCTACCTGACAGATAATCAAGCTTACAGTCATAT 165

Qy 21 TyrIleLeuLys 24

Db 166 CATCTGCTAAAG 177

RESULT 7

AX462181
LOCUS AX462181 498 bp DNA linear PAT 09-JUL-2002
DEFINITION Sequence 5 from Patent WO0238599.
ACCESSION AX462181
VERSION AX462181.1 GI:21727744
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1 Scheres,B.J., Bllou,I. and Folmer,S.D.
AUTHORS A plant development regulating gene and its uses
TITLE Patent: WO 0238599-A 5 16-MAY-2002;
JOURNAL Universiteit Utrecht (NL)

FEATURES

Location/Qualifiers
source
1..498
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

BASE COUNT 136 a 109 c 105 g 148 t
ORIGIN

Alignment Scores:

Pred. No.: 4.83e-06 Length: 498
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3

Best Local Similarity: 75.00% Mismatches: 3
 Query Match: 75.21% Indels: 0
 DB: 6 Gaps: 0

US-10-036-492-6 (1-24) x AX462181 (1-498)

QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 DB 106 GTTAATTGCGAGCTATTAGCCACCAGCTACCTGCAGATAATCAAGCTTACAGTGCATAT 165

QY 21 TyrIleLeuLys 24
 DB 166 CATCTGCTAAAG 177

RESULT 8
 AX462198
 LOCUS
 DEFINITION Sequence 22 from Patent WO0238599.
 ACCESSION AX462198
 VERSION AX462198.1 GI:21727757
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE
 1 Scheres,B.J., Bllou,I. and Folmer,S.D.
 A plant development regulating gene and its uses
 Patent: WO 0238599-A 22 16-MAY-2002;
 Universiteit Utrecht (NL)

FEATURES
 source
 1..1557
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"

BASE COUNT 448 a 339 c 328 g 442 t

ORIGIN
 Alignment Scores:
 Pred. No.: 1.59e-05 Length: 1557
 Score: 91.00 Matches: 18
 Percent Similarity: 87.50% Conservative: 3
 Best Local Similarity: 75.00% Mismatches: 3
 Query Match: 75.21% Indels: 0
 DB: 6 Gaps: 0

US-10-036-492-6 (1-24) x AX462198 (1-1557)

QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 DB 106 GTTAATTGCGAGCTATTAGCCACCAGCTACCTGCAGATAATCAAGCTTACAGTGCATAT 165

QY 21 TyrIleLeuLys 24
 DB 166 CATCTGCTAAAG 177

RESULT 9
 AY128780
 LOCUS
 DEFINITION Arabidopsis thaliana CDC27/NUC2-like protein (At2g20000) mRNA,
 complete cds.
 ACCESSION AY128780
 VERSION AY128780.1 GI:22136203
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE
 1 (bases 1 to 1915)
 Tripp,M., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M.,
 Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,

Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
 Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
 Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
 Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
 Submitted (01-JUL-2002) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 e-mail for correspondence: arab@sequence.stanford.edu

TITLE
 JOURNAL

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN
 Arabidopsis Full-length cDNA"): Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members constructed and
 sequenced the PUNI (ORF) clones using the RAFL cDNAs: Tripp,M.,
 Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M.,
 Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R.,
 Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H.,
 Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Tripp,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally
 to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
 (SSP/Stanford) contributed equally to this work as PIs.

FEATURES
 Location/Qualifiers
 1..1915
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosomes="2"
 /clone="U16204"
 /note="This clone is in PENTR/SD-dTopo This is a cloned
 PCR product using RIKEN clone RAFL07-12-N15 (AY062470) as
 a template
 ecotype: Columbia"
 1..1915
 /gene="At2g20000"
 /note="synonym: T2G17.20"
 1..1416
 /gene="At2g20000"
 /codon_start=1
 /product="CDC27/NUC2-like protein"
 /protein_id="AA091180.1"
 /db_xref="GI:22136204"
 /translation="MEAMLVDCVNNSLRHFVYKAI FMCERLC AEPSEVNQLLATS
 YLQNOAYSA YHLKGTQMAQRYL FALSCFQMDLNEAESALCPYNEPQAI PINGAA
 GHYLLGLIYKYTD RKNAAQFKQSLTIDPLLWAAVEELCILGAEEATAVFGTAAL
 SIQKVMQQLSTSLGNTVNEERNSTSTKNTSSDYSPQSKHTQSHGLKDISGNFHS
 HGNGVGNMSPYNTSPVAQLSGIAPPLRPNPQPAVANPNSLITDSKPKSTVNST
 LQAPRRKFDEGKLKIKISRLRSDSGPRSSRLSDSGANINSSVATVSGNVNASKY
 LGGSKLUSLALRSVTLRKHSWANENMDEGVGEPPDDSRPNTASTTIGSMASNDQDE
 TWSIGGIAMSSQTTITGVSEILNLLRTLGGCRSLSYMYRCQALDITMKLPHRYNTG
 WLSQVTSDFSFRFLPYMDIA"

BASE COUNT 558 a 405 c 404 g 548 t

ORIGIN
 Alignment Scores:
 Pred. No.: 1.97e-05 Length: 1915
 Score: 91.00 Matches: 18
 Percent Similarity: 87.50% Conservative: 3
 Best Local Similarity: 75.00% Mismatches: 3
 Query Match: 75.21% Indels: 0
 DB: 8 Gaps: 0

US-10-036-492-6 (1-24) x AY128780 (1-1915)

QY 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 DB 106 GTTAATTGCGAGCTATTAGCCACCAGCTACCTGCAGATAATCAAGCTTACAGTGCATAT 165

QY 21 TyrIleLeuLys 24

```

Db      166 CATCTGCTAAAG 177
      :::::|||||
RESULT 10
AX462197          2115 bp          DNA          linear          PAT 09-JUL-2002
LOCUS
DEFINITION
Sequence 21 from Patent WO0238599.
ACCESSION
AX462197
VERSION
AX462197.1 GI:21727756
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
1 Scheres,B.J., Bllou,I. and Folmer,S.D.
AUTHORS
A plant development regulating gene and its uses
TITLE
Patent: WO 0238599-A 21 16-MAY-2002;
JOURNAL
Universiteit Utrecht (NL)
FEATURES
Location/Qualifiers
source
1..2115
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT
634 a 450 c 458 g 573 t
ORIGIN
Alignment Scores:
Pred. No.: 2,19e-05 Length: 2115
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 6 Gaps: 0

US-10-036-492-6 (1-24) x AX462197 (1-2115)
Qy      1 ValAsnLeuGlnLeuAlaAargCystYrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db      106 GTTAATTTCAGCTATTAGCCAGCTACCTGCGAGATAATCAAGCTTACAGTGCATAT 165
      :::::|||||
Qy      21 TyrIleLeuLys 24
Db      166 CATCTGCTAAAG 177
      :::::|||||
RESULT 11
AX462196          2169 bp          DNA          linear          PAT 09-JUL-2002
LOCUS
DEFINITION
Sequence 20 from Patent WO0238599.
ACCESSION
AX462196
VERSION
AX462196.1 GI:21727755
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
1 Scheres,B.J., Bllou,I. and Folmer,S.D.
AUTHORS
A plant development regulating gene and its uses
TITLE
Patent: WO 0238599-A 20 16-MAY-2002;
JOURNAL
Universiteit Utrecht (NL)
FEATURES
Location/Qualifiers
source
1..2169
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT
643 a 462 c 470 g 594 t
ORIGIN
Alignment Scores:
Pred. No.: 2,24e-05 Length: 2169
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 6 Gaps: 0

US-10-036-492-6 (1-24) x AX462196 (1-2169)
Qy      1 ValAsnLeuGlnLeuAlaAargCystYrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db      106 GTTAATTTCAGCTATTAGCCAGCTACCTGCGAGATAATCAAGCTTACAGTGCATAT 165
      :::::|||||
Qy      21 TyrIleLeuLys 24
Db      166 CATCTGCTAAAG 177
      :::::|||||
RESULT 13
AX462179          2235 bp          DNA          linear          PAT 09-JUL-2002
LOCUS
DEFINITION
Sequence 3 from Patent WO0238599.
ACCESSION
AX462179
VERSION
AX462179.1 GI:21727742
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
1 Scheres,B.J., Bllou,I. and Folmer,S.D.
AUTHORS

```

```

Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 6 Gaps: 0

US-10-036-492-6 (1-24) x AX462196 (1-2169)
Qy      1 ValAsnLeuGlnLeuAlaAargCystYrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db      106 GTTAATTTCAGCTATTAGCCAGCTACCTGCGAGATAATCAAGCTTACAGTGCATAT 165
      :::::|||||
Qy      21 TyrIleLeuLys 24
Db      166 CATCTGCTAAAG 177
      :::::|||||
RESULT 12
AX068992          2220 bp          DNA          linear          PAT 25-JAN-2001
LOCUS
DEFINITION
Sequence 15 from Patent WO0102430.
ACCESSION
AX068992
VERSION
AX068992.1 GI:12578822
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
1 Hemerly,A.S., Ferreira,P.C. and Rombauts,S.
AUTHORS
Arabidopsis thaliana cdc7 and cdc27 homologs
TITLE
Patent: WO 0102430-A 15 11-JAN-2001;
JOURNAL
CropDesign N.V. (BE) ; Universidade Federal do Rio de Janeiro (BR)
FEATURES
Location/Qualifiers
source
1..2220
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT
658 a 474 c 480 g 608 t
ORIGIN
Alignment Scores:
Pred. No.: 2,3e-05 Length: 2220
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 6 Gaps: 0

US-10-036-492-6 (1-24) x AX068992 (1-2220)
Qy      1 ValAsnLeuGlnLeuAlaAargCystYrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db      106 GTTAATTTCAGCTATTAGCCAGCTACCTGCGAGATAATCAAGCTTACAGTGCATAT 165
      :::::|||||
Qy      21 TyrIleLeuLys 24
Db      166 CATCTGCTAAAG 177
      :::::|||||
RESULT 13
AX462179          2235 bp          DNA          linear          PAT 09-JUL-2002
LOCUS
DEFINITION
Sequence 3 from Patent WO0238599.
ACCESSION
AX462179
VERSION
AX462179.1 GI:21727742
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
1 Scheres,B.J., Bllou,I. and Folmer,S.D.
AUTHORS

```

TITLE A plant development regulating gene and its uses

JOURNAL Patent: WO 0238599-A 3 16-MAY-2002;

UNIVERSITEIT UTRECHT (NL)

FEATURES Location/Qualifiers

source

1. .2235

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

BASE COUNT 664 a 475 c 484 g 612 t

ORIGIN

Alignment Scores:

Pred. No.: 2.32e-05 Length: 2235
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 6 Gaps: 0

US-10-036-492-6 (1-24) x AX462179 (1-2235)

QY 1 ValAsnLeuGlnLeuLeuAlaAArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
|||||
DB 106 GTTAATTGGAGCTATTAGCCAGCTACCTCGAGAAATATCAAGCTTACAGTGCATAT 165

QY 21 TyrIleLeuLys 24

DB 166 CATCTGCTAAAG 177

RESULT 14

AX462193

LOCUS

DEFINITION

AX462193

ACCESSION

VERSION

AX462193.1

GI:21727752

KEYWORDS

SOURCE

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 Scheres B.J., Bilou, I. and Folmer, S.D.

A plant development regulating gene and its uses

Patent: WO 0238599-A 17 16-MAY-2002;

UNIVERSITEIT UTRECHT (NL)

FEATURES

source

1. .2235

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

BASE COUNT 664 a 474 c 484 g 613 t

ORIGIN

Alignment Scores:

Pred. No.: 2.32e-05 Length: 2235
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 6 Gaps: 0

US-10-036-492-6 (1-24) x AX462193 (1-2235)

QY 1 ValAsnLeuGlnLeuLeuAlaAArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
|||||
DB 106 GTTAATTGGAGCTATTAGCCAGCTACCTCGAGAAATATCAAGCTTACAGTGCATAT 165

QY 21 TyrIleLeuLys 24

DB 166 CATCTGCTAAAG 177

RESULT 15

AX462195

LOCUS

DEFINITION

AX462195

ACCESSION

VERSION

AX462195.1

GI:21727754

KEYWORDS

SOURCE

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 Scheres B.J., Bilou, I. and Folmer, S.D.

A plant development regulating gene and its uses

Patent: WO 0238599-A 19 16-MAY-2002;

UNIVERSITEIT UTRECHT (NL)

FEATURES

source

1. .2235

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

BASE COUNT 665 a 475 c 483 g 612 t

ORIGIN

Alignment Scores:

Pred. No.: 2.32e-05 Length: 2235
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 75.21% Indels: 0
DB: 6 Gaps: 0

US-10-036-492-6 (1-24) x AX462195 (1-2235)

QY 1 ValAsnLeuGlnLeuLeuAlaAArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
|||||
DB 106 GTTAATTGGAGCTATTAGCCAGCTACCTCGAGAAATATCAAGCTTACAGTGCATAT 165

QY 21 TyrIleLeuLys 24

DB 166 CATCTGCTAAAG 177

Search completed: February 2, 2004, 15:40:00

Job time : 6296 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 2, 2004, 12:47:18 ; Search time 247 Seconds
(without alignments)
20.193 Million cell updates/sec

Title: US-10-036-492-6

Perfect score: 121

Sequence: 1 VNLQLLARCYLNSQAYSAYILK 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	121	100.0	24	14	US-10-036-492-6
2	121	100.0	716	14	US-10-036-492-11
3	121	100.0	728	14	US-10-036-492-5
4	91	75.2	161	14	US-10-036-492-10
5	91	75.2	739	14	US-10-036-492-13
6	46	38.0	556	9	US-09-887-586A-32
7	46	38.0	556	9	US-09-903-012-32
8	46	38.0	556	11	US-09-900-797-32
9	46	38.0	882	12	US-10-104-047-3543
10	44.5	36.8	167	12	US-10-191-813-65
11	44.5	36.8	205	12	US-10-191-813-55
12	44.5	36.8	205	12	US-10-191-813-59
13	44.5	36.8	205	12	US-10-191-813-61
14	44.5	36.8	205	12	US-10-191-813-63
15	44.5	36.8	205	12	US-10-191-813-68

Sequence 70, Appl
Sequence 120, Appl
Sequence 122, Appl
Sequence 7, Appl
Sequence 1496, Appl
Sequence 14, Appl
Sequence 391, Appl
Sequence 391, Appl
Sequence 82, Appl
Sequence 342, Appl
Sequence 236, Appl
Sequence 74, Appl
Sequence 306, Appl
Sequence 33823, A
Sequence 5799, Ap
Sequence 2904, Ap
Sequence 1241, Ap
Sequence 7660, Ap
Sequence 10, Appl
Sequence 65, Appl
Sequence 65, Appl
Sequence 6050, Ap
Sequence 6049, Ap
Sequence 6048, Ap
Sequence 1848, Ap
Sequence 603, Appl
Sequence 3264, Ap
Sequence 4668, Ap
Sequence 72, Appl

US-10-191-813-70
US-10-191-813-120
US-10-191-813-122
US-10-361-522-7
US-09-867-550-1496
US-09-915-043-14
US-09-915-043-16
US-09-886-055-391
US-09-804-291-391
US-10-017-161-82
US-10-017-161-342
US-10-387-629-236
US-10-292-798-74
US-10-292-798-306
US-09-864-761-33823
US-10-369-493-5799
US-10-104-047-2904
US-10-033-357-1241
US-10-032-585-7660
US-10-134-102-10
US-10-126-927-65
US-10-126-931A-65
US-10-369-493-6050
US-10-369-493-6049
US-10-369-493-6048
US-09-864-408A-1848
US-09-833-245-609
US-10-320-797-3264
US-09-738-626-4668
US-10-387-629-72

16 44.5 36.8 205 12
17 44.5 36.8 797 12
18 44.5 36.8 797 12
19 44.5 36.8 1771 12
20 44 36.4 273 9
21 44 36.4 311 11
22 44 36.4 311 11
23 44 36.4 312 10
24 44 36.4 312 11
25 44 36.4 312 12
26 44 36.4 312 12
27 44 36.4 312 12
28 44 36.4 312 12
29 44 36.4 312 12
30 44 36.4 484 9
31 44 36.4 504 12
32 44 36.4 638 12
33 43.5 36.0 77 15
34 43 35.5 448 12
35 43 35.5 513 12
36 43 35.5 932 15
37 43 35.5 932 15
38 43 35.5 1851 12
39 43 35.5 1966 12
40 43 35.5 2848 12
41 42 34.7 58 12
42 42 34.7 318 12
43 42 34.7 597 12
44 41 33.9 221 10
45 41 33.9 304 12

ALIGNMENTS

RESULT 1
US-10-036-492-6
; Sequence 6, Application US/10036492
; Publication No. US20020164757A1
; GENERAL INFORMATION:
; APPLICANT: HEMERLY, ADRIANA
; APPLICANT: FERREIRA, PAULO
; APPLICANT: ROMBAUTS, STEPHANE
; TITLE OF INVENTION: PLANT DNA REPLICATION MODULATING PROTEINS
; FILE REFERENCE: 217943USOXCONT
; CURRENT APPLICATION NUMBER: US/10/036,492
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: EP 99202214.5
; PRIOR FILING DATE: 1999-07-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-036-492-6

Query Match 100.0%; Score 121; DB 14; Length 24;
Best Local Similarity 100.0%; Pred. No. 6.9e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNLQLLARCYLNSQAYSAYILK 24
DB 1 VNLQLLARCYLNSQAYSAYILK 24

RESULT 2
US-10-036-492-11
; Sequence 11, Application US/10036492
; Publication No. US20020164757A1
; GENERAL INFORMATION:

```
; APPLICANT: HEMERLY, ADRIANA
; APPLICANT: FERREIRA, PAULO
; APPLICANT: ROMBAUTS, STEPHANE
; TITLE OF INVENTION: PLANT DNA REPLICATION MODULATING PROTEINS
; FILE REFERENCE: 217943USOXCONT
; CURRENT APPLICATION NUMBER: US/10/036,492
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: EP 99202214.5
; PRIOR FILING DATE: 1999-07-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 716
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-036-492-11

Query Match          100.0%; Score 121; DB 14; Length 716;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNQLLARCYLNSQAYSAYIYLK 24
DB 37 VNQLLARCYLNSQAYSAYIYLK 60

RESULT 3
US-10-036-492-5
; Sequence 5, Application US/10036492
; Publication No. US20020164757A1
; GENERAL INFORMATION:
; APPLICANT: HEMERLY, ADRIANA
; APPLICANT: FERREIRA, PAULO
; APPLICANT: ROMBAUTS, STEPHANE
; TITLE OF INVENTION: PLANT DNA REPLICATION MODULATING PROTEINS
; FILE REFERENCE: 217943USOXCONT
; CURRENT APPLICATION NUMBER: US/10/036,492
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: EP 99202214.5
; PRIOR FILING DATE: 1999-07-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-036-492-5

Query Match          100.0%; Score 121; DB 14; Length 728;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNQLLARCYLNSQAYSAYIYLK 24
DB 37 VNQLLARCYLNSQAYSAYIYLK 60

RESULT 4
US-10-036-492-10
; Sequence 10, Application US/10036492
; Publication No. US20020164757A1
; GENERAL INFORMATION:
; APPLICANT: HEMERLY, ADRIANA
; APPLICANT: FERREIRA, PAULO
; APPLICANT: ROMBAUTS, STEPHANE
; TITLE OF INVENTION: PLANT DNA REPLICATION MODULATING PROTEINS
; FILE REFERENCE: 217943USOXCONT
; CURRENT APPLICATION NUMBER: US/10/036,492
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: EP 99202214.5
; PRIOR FILING DATE: 1999-07-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-036-492-10

Query Match          75.2%; Score 91; DB 14; Length 161;
Best Local Similarity 75.0%; Pred. No. 2.3e-06;
Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VNQLLARCYLNSQAYSAYIYLK 24
DB 36 VNQLLATSYLQNNQAYSAYHLK 59

RESULT 5
US-10-036-492-13
; Sequence 13, Application US/10036492
; Publication No. US20020164757A1
; GENERAL INFORMATION:
; APPLICANT: HEMERLY, ADRIANA
; APPLICANT: FERREIRA, PAULO
; APPLICANT: ROMBAUTS, STEPHANE
; TITLE OF INVENTION: PLANT DNA REPLICATION MODULATING PROTEINS
; FILE REFERENCE: 217943USOXCONT
; CURRENT APPLICATION NUMBER: US/10/036,492
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: EP 99202214.5
; PRIOR FILING DATE: 1999-07-05
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-036-492-13

Query Match          75.2%; Score 91; DB 14; Length 739;
Best Local Similarity 75.0%; Pred. No. 1.2e-05;
Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VNQLLARCYLNSQAYSAYIYLK 24
DB 36 VNQLLATSYLQNNQAYSAYHLK 59

RESULT 6
US-09-887-586A-32
; Sequence 32, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: NO. US20020094556A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
```

```
US-09-887-586A-32
Query Match      38.0%; Score 46; DB 9; Length 556;
Best Local Similarity 64.3%; Pred. No. 81;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      10 YLSNSQAYSAYYIL 23
      ||||:||||:|
Db      402 YLSNALATSTYYLL 415

RESULT 7
US-09-903-012-32
; Sequence 32, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-903-012-32

Query Match      38.0%; Score 46; DB 9; Length 556;
Best Local Similarity 64.3%; Pred. No. 81;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      10 YLSNSQAYSAYYIL 23
      ||||:||||:|
Db      402 YLSNALATSTYYLL 415

US-09-900-797-32
Query Match      38.0%; Score 46; DB 11; Length 556;
Best Local Similarity 64.3%; Pred. No. 81;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      10 YLSNSQAYSAYYIL 23
      ||||:||||:|
Db      402 YLSNALATSTYYLL 415

RESULT 9
US-10-104-047-3543
; Sequence 3543, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1 full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3543
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3543

Query Match      38.0%; Score 46; DB 12; Length 882;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY      3 LQLLARCVLSNSQAYSAY 20
      |::|::|::|::|
Db      783 LELLRCVQYNKSCYKAY 800

RESULT 10
US-10-191-813-65
; Sequence 65, Application US/10191813
; Publication No. US20030165907A1
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/10/191,813
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US/09/191,468
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 65
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Human
US-10-191-813-65

Query Match      36.8%; Score 44.5; DB 12; Length 167;
Best Local Similarity 40.0%; Pred. No. 38;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY      1 VNQLLLA-RCVLSNSQAYSAYYILK 24
      ||||:||||:|
Db      131 VNFTSLANKTFVSGSEYFKYFLVK 155

RESULT 11
```



```
US-10-191-813-55
; Sequence 55, Application US/10191813
; Publication No. US20030165907A1
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/10/191,813
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US/09/191,468
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Human
US-10-191-813-55

Query Match          36.8%; Score 44.5; DB 12; Length 205;
Best Local Similarity 40.0%; Pred. No. 47;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY      1 VNQLLLA-RCYLSNSQAYSAYILK 24
      ||| : : : : : : : : : : : :
Db      131 VNFTSLANKTFVSGSEYFKFVLK 155

RESULT 12
US-10-191-813-59
; Sequence 59, Application US/10191813
; Publication No. US20030165907A1
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/10/191,813
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US/09/191,468
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Human
US-10-191-813-59

Query Match          36.8%; Score 44.5; DB 12; Length 205;
Best Local Similarity 40.0%; Pred. No. 47;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY      1 VNQLLLA-RCYLSNSQAYSAYILK 24
      ||| : : : : : : : : : : : :
Db      131 VNFTSLANKTFVSGSEYFKFVLK 155

RESULT 13
US-10-191-813-61
; Sequence 61, Application US/10191813
; Publication No. US20030165907A1
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/10/191,813
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US/09/191,468
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Human
US-10-191-813-61

Query Match          36.8%; Score 44.5; DB 12; Length 205;
Best Local Similarity 40.0%; Pred. No. 47;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY      1 VNQLLLA-RCYLSNSQAYSAYILK 24
      ||| : : : : : : : : : : : :
Db      131 VNFTSLANKTFVSGSEYFKFVLK 155

RESULT 14
US-10-191-813-63
; Sequence 63, Application US/10191813
; Publication No. US20030165907A1
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/10/191,813
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US/09/191,468
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Human
US-10-191-813-63

Query Match          36.8%; Score 44.5; DB 12; Length 205;
Best Local Similarity 40.0%; Pred. No. 47;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY      1 VNQLLLA-RCYLSNSQAYSAYILK 24
      ||| : : : : : : : : : : : :
Db      131 VNFTSLANKTFVSGSEYFKFVLK 155

RESULT 15
US-10-191-813-68
; Sequence 68, Application US/10191813
; Publication No. US20030165907A1
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/10/191,813
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US/09/191,468
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Human
US-10-191-813-68

Query Match          36.8%; Score 44.5; DB 12; Length 205;
```

Best Local Similarity 40.0%; Pred. No. 47;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 VNLQLLA-RCYLSNSQAYSYYILK 24
|| : : : : : : : : : : : : : : : :
Db 131 VNFTSLANKTFVSGSEBEYFKYFVLK 155

Search completed: February 2, 2004, 13:25:00
Job time : 247 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 2, 2004, 12:50:28 ; Search time 76 Seconds
(without alignments)
13.361 Million cell updates/sec

Title: US-10-036-492-6
Perfect score: 121
Sequence: 1 VNLQLARCYLSNQAYSAYYILK 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	38.0	213	4 US-09-134-001C-5360	Sequence 5360, Ap
2	46	38.0	556	4 US-09-398-395A-32	Sequence 32, Appl
3	46	38.0	556	4 US-09-887-586A-32	Sequence 32, Appl
4	46	38.0	556	4 US-09-895-752-32	Sequence 32, Appl
5	46	38.0	556	4 US-09-903-012B-32	Sequence 32, Appl
6	44.5	36.8	167	4 US-09-191-468-65	Sequence 65, Appl
7	44.5	36.8	205	4 US-09-191-468-55	Sequence 55, Appl
8	44.5	36.8	205	4 US-09-191-468-59	Sequence 59, Appl
9	44.5	36.8	205	4 US-09-191-468-61	Sequence 61, Appl
10	44.5	36.8	205	4 US-09-191-468-63	Sequence 63, Appl
11	44.5	36.8	205	4 US-09-191-468-68	Sequence 68, Appl
12	44.5	36.8	205	4 US-09-191-468-70	Sequence 70, Appl
13	44.5	36.8	797	4 US-09-191-468-120	Sequence 120, App
14	44.5	36.8	797	4 US-09-191-468-122	Sequence 122, App
15	44.5	36.8	831	2 US-09-047-026A-4	Sequence 4, Appl
16	44	36.4	308	4 US-09-252-991A-28210	Sequence 28210, A
17	43	35.5	299	2 US-09-047-026A-24	Sequence 24, Appl
18	42	34.7	634	4 US-09-328-352-4627	Sequence 4627, Ap
19	41	33.9	150	4 US-09-107-532A-5128	Sequence 5128, Ap
20	41	33.9	161	4 US-09-107-532A-7186	Sequence 7186, Ap
21	40	33.1	205	4 US-09-191-468-57	Sequence 57, Appl
22	40	33.1	354	2 US-08-700-013B-9	Sequence 9, Appl
23	40	33.1	405	2 US-08-700-013B-6	Sequence 6, Appl
24	40	33.1	434	2 US-08-484-575A-20	Sequence 20, Appl
25	40	33.1	434	3 US-08-477-459-20	Sequence 20, Appl
26	40	33.1	434	3 US-08-479-869-20	Sequence 20, Appl
27	40	33.1	434	3 US-08-486-414-46	Sequence 46, Appl

28 40 33.1 434 5 PCT-US94-01826A-20 Sequence 20, Appl
29 40 33.1 434 5 PCT-US94-02252A-20 Sequence 20, Appl
30 40 33.1 434 5 PCT-US96-03916-11 Sequence 11, Appl
31 40 33.1 434 5 PCT-US96-03916-67 Sequence 67, Appl
32 40 33.1 548 4 US-09-398-395A-2 Sequence 2, Appl
33 40 33.1 548 4 US-09-398-395A-4 Sequence 4, Appl
34 40 33.1 548 4 US-09-398-395A-6 Sequence 6, Appl
35 40 33.1 548 4 US-09-398-395A-8 Sequence 8, Appl
36 40 33.1 548 4 US-09-398-395A-10 Sequence 10, Appl
37 40 33.1 548 4 US-09-887-586A-2 Sequence 2, Appl
38 40 33.1 548 4 US-09-887-586A-4 Sequence 4, Appl
39 40 33.1 548 4 US-09-887-586A-6 Sequence 6, Appl
40 40 33.1 548 4 US-09-887-586A-8 Sequence 8, Appl
41 40 33.1 548 4 US-09-887-586A-10 Sequence 10, Appl
42 40 33.1 548 4 US-09-895-752-2 Sequence 2, Appl
43 40 33.1 548 4 US-09-895-752-4 Sequence 4, Appl
44 40 33.1 548 4 US-09-895-752-6 Sequence 6, Appl
45 40 33.1 548 4 US-09-895-752-8 Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-09-134-001C-5360
; Sequence 5360, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5360
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5360

Query Match 38.0%; Score 46; DB 4; Length 213;
Best Local Similarity 69.2%; Pred No. 6.4; Mismatches 4; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 8 RCYLSNQAYSAY 20

DB 89 RCYNSNYDFYSAY 101

RESULT 2

US-09-398-395A-32
; Sequence 32, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6468772, Joseph P.
; APPLICANT: Starks, Courtney R.
; APPLICANT: Hanna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23

; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-398-395A-32

Query Match 38.0%; Score 46; DB 4; Length 556;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 YLSNSQAYSAYYIL 23
||||:|:|:|:|:
Db 402 YLSNALATSTYILL 415

RESULT 3

US-09-887-586A-32
; Sequence 32, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6495354, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-887-586A-32

Query Match 38.0%; Score 46; DB 4; Length 556;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 YLSNSQAYSAYYIL 23
||||:|:|:|:|:
Db 402 YLSNALATSTYILL 415

RESULT 4

US-09-895-752-32
; Sequence 32, Application US/09895752
; Patent No. 6559297
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6559297, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/895,752
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262

; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-895-752-32

Query Match 38.0%; Score 46; DB 4; Length 556;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 YLSNSQAYSAYYIL 23
||||:|:|:|:|:
Db 402 YLSNALATSTYILL 415

RESULT 5

US-09-903-012B-32
; Sequence 32, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6569656, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-903-012B-32

Query Match 38.0%; Score 46; DB 4; Length 556;
Best Local Similarity 64.3%; Pred. No. 20;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 YLSNSQAYSAYYIL 23
||||:|:~|:|:|:|:
Db 402 YLSNALATSTYILL 415

RESULT 6

US-09-191-468-65
; Sequence 65, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/09/191,468A
; CURRENT FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 65
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Human
US-09-191-468-65

Query Match 36.8%; Score 44.5; DB 4; Length 167;
Best Local Similarity 40.0%; Pred. No. 8.4;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 VNQLLA-RCVLSNSQAYSAYVILK 24
DB 131 VNFTSLANKTFVSGSEYFKYFVLK 155

RESULT 7

US-09-191-468-55
; Sequence 55, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/09/191,468A
; CURRENT FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Human
US-09-191-468-55

Query Match 36.8%; Score 44.5; DB 4; Length 205;
Best Local Similarity 40.0%; Pred. No. 11;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 VNQLLA-RCVLSNSQAYSAYVILK 24
DB 131 VNFTSLANKTFVSGSEYFKYFVLK 155

RESULT 8

US-09-191-468-59
; Sequence 59, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/09/191,468A
; CURRENT FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Human
US-09-191-468-59

Query Match 36.8%; Score 44.5; DB 4; Length 205;
Best Local Similarity 40.0%; Pred. No. 11;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 VNQLLA-RCVLSNSQAYSAYVILK 24
DB 131 VNFTSLANKTFVSGSEYFKYFVLK 155

RESULT 9

US-09-191-468-61
; Sequence 61, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:

; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/09/191,468A
; CURRENT FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Human
US-09-191-468-61

Query Match 36.8%; Score 44.5; DB 4; Length 205;
Best Local Similarity 40.0%; Pred. No. 11;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 VNQLLA-RCVLSNSQAYSAYVILK 24
DB 131 VNFTSLANKTFVSGSEYFKYFVLK 155

RESULT 10

US-09-191-468-63
; Sequence 63, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/09/191,468A
; CURRENT FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Human
US-09-191-468-63

Query Match 36.8%; Score 44.5; DB 4; Length 205;
Best Local Similarity 40.0%; Pred. No. 11;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 1 VNQLLA-RCVLSNSQAYSAYVILK 24
DB 131 VNFTSLANKTFVSGSEYFKYFVLK 155

RESULT 11

US-09-191-468-68
; Sequence 68, Application US/09191468A
; Patent No. 6416975
; GENERAL INFORMATION:
; APPLICANT: Gallagher, Michael J.
; APPLICANT: Burgess, Loyd R.
; APPLICANT: Brunden, Kurt R.
; TITLE OF INVENTION: Human Glycine Transporter Type 2
; FILE REFERENCE: 12311US01
; CURRENT APPLICATION NUMBER: US/09/191,468A
; CURRENT FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Human
US-09-191-468-68

Query Match 36.8%; Score 44.5; DB 4; Length 205;
Best Local Similarity 40.0%; Pred. No. 11;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
QY 1 VNQLLA-RCYLSNSQAYSAYILK 24
DB 131 VNFTSLANKTFVSGSEYFKYFLK 155
RESULT 12
US-09-191-468-70
Sequence 70, Application US/09191468A
Patent No. 6416975
GENERAL INFORMATION:
APPLICANT: Gallagher, Michael J.
APPLICANT: Burgess, Loyd R.
APPLICANT: Brunden, Kurt R.
TITLE OF INVENTION: Human Glycine Transporter Type 2
FILE REFERENCE: 12311US01
CURRENT APPLICATION NUMBER: US/09/191,468A
CURRENT FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 124
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 70
LENGTH: 205
TYPE: PRT
ORGANISM: Human
US-09-191-468-70
Query Match 36.8%; Score 44.5; DB 4; Length 205;
Best Local Similarity 40.0%; Pred. No. 11;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
QY 1 VNQLLA-RCYLSNSQAYSAYILK 24
DB 131 VNFTSLANKTFVSGSEYFKYFLK 155
RESULT 13
US-09-191-468-120
Sequence 120, Application US/09191468A
Patent No. 6416975
GENERAL INFORMATION:
APPLICANT: Gallagher, Michael J.
APPLICANT: Burgess, Loyd R.
APPLICANT: Brunden, Kurt R.
TITLE OF INVENTION: Human Glycine Transporter Type 2
FILE REFERENCE: 12311US01
CURRENT APPLICATION NUMBER: US/09/191,468A
CURRENT FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 124
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 120
LENGTH: 797
TYPE: PRT
ORGANISM: Human
US-09-191-468-120
Query Match 36.8%; Score 44.5; DB 4; Length 797;
Best Local Similarity 40.0%; Pred. No. 56;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
QY 1 VNQLLA-RCYLSNSQAYSAYILK 24
DB 357 VNFTSLANKTFVSGSEYFKYFLK 381
RESULT 14
US-09-191-468-122
Sequence 122, Application US/09191468A
Patent No. 6416975
GENERAL INFORMATION:
APPLICANT: Gallagher, Michael J.

APPLICANT: Burgess, Loyd R.
APPLICANT: Brunden, Kurt R.
TITLE OF INVENTION: Human Glycine Transporter Type 2
FILE REFERENCE: 12311US01
CURRENT APPLICATION NUMBER: US/09/191,468A
CURRENT FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 124
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 122
LENGTH: 797
TYPE: PRT
ORGANISM: Human
US-09-191-468-122
Query Match 36.8%; Score 44.5; DB 4; Length 797;
Best Local Similarity 40.0%; Pred. No. 56;
Matches 10; Conservative 6; Mismatches 8; Indels 1; Gaps 1;
QY 1 VNQLLA-RCYLSNSQAYSAYILK 24
DB 357 VNFTSLANKTFVSGSEYFKYFLK 381
RESULT 15
US-09-047-026A-4
Sequence 4, Application US/09047026A
Patent No. 5989897
GENERAL INFORMATION:
APPLICANT: Pillus, Lorraine
APPLICANT: Clarke, Astrid
APPLICANT: Lowell, Joanna
APPLICANT: Jacobson, Sandra
APPLICANT: Reifsnnyder, Cheryl
TITLE OF INVENTION: Yeast Silencing Genes, Proteins and
TITLE OF INVENTION: Methods
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/047,026A
FILING DATE: 24-MAR-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/042,375
FILING DATE: 24-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 1-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 831 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-047-026A-4
Query Match 36.8%; Score 44.5; DB 2; Length 831;
Best Local Similarity 37.0%; Pred. No. 59;
Matches 10; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

QY 2 NLQLLARCYLSNSQAY-----SNYYIL 23
|| |||:|:::|
Db 355 NLCLAKCFINSKTYDYDVEPFIYIL 381

Search completed: February 2, 2004, 13:26:27
Job time : 76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 2, 2004, 12:08:59 ; Search time 648 Seconds
(without alignments)
5.879 Million cell updates/sec

Title: US-10-036-492-6
Perfect score: 121
Sequence: 1 VNLQLLARCYSLSQAYSAYVILK 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues 1107863
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq 19Jun03.*
1: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-embl/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	121	100.0	24	22	AA668953	Arabidopsis thalia
2	121	100.0	716	22	AA668956	Arabidopsis thalia
3	121	100.0	728	22	AA668952	Arabidopsis thalia
4	106.5	88.0	23	22	AA668960	Arabidopsis thalia
5	91	75.2	161	22	AA668955	Arabidopsis thalia
6	91	75.2	162	23	ABG65520	Cell cycle regulat
7	91	75.2	166	23	ABG65521	Cell cycle regulat
8	91	75.2	518	23	ABG65529	Functionally inact
9	91	75.2	704	23	ABG65528	Functionally inact

10	91	75.2	722	23	ABG65527	Functionally inact
11	91	75.2	739	22	AA668958	Arabidopsis thalia
12	91	75.2	740	23	ABG65522	Cell cycle regulat
13	91	75.2	743	23	ABG65524	Functionally inact
14	91	75.2	744	23	ABG65523	Cell cycle regulat
15	91	75.2	744	23	ABG65526	Functionally inact
16	91	75.2	744	23	ABG65530	Functionally inact
17	91	75.2	748	21	AA650614	Arabidopsis thalia
18	91	75.2	768	21	AA650613	Functionally inact
19	91	75.2	770	23	ABG65525	Arabidopsis thalia
20	91	75.2	771	21	AA650612	Arabidopsis thalia
21	59	48.8	62	21	AA601804	Human secreted pro
22	59	48.8	823	17	AAW00364	Human CDC27. Homo
23	59	48.8	824	16	AA675848	H-NUC retinoblasto
24	59	48.8	881	22	ABG05595	Novel human diagno
25	49	40.5	900	22	ABG06088	Drosophila melanog
26	46	38.0	213	23	ABP40515	Staphylococcus epi
27	46	38.0	213	23	ABP40515	Potato vetaspiradi
28	46	38.0	556	21	AA690843	Drosophila melanog
29	45	37.2	887	22	ABG68631	Human ORFX protein
30	44.5	36.8	92	23	ABP05350	Human glycine tran
31	44.5	36.8	167	21	AA613518	Human glycine tran
32	44.5	36.8	205	21	AA609892	Human glycine tran
33	44.5	36.8	205	21	AA621165	Human glycine tran
34	44.5	36.8	205	21	AA621166	Human glycine tran
35	44.5	36.8	205	21	AA621167	Human glycine tran
36	44.5	36.8	205	21	AA621168	Human glycine tran
37	44.5	36.8	205	21	AA621169	Human glycine tran
38	44.5	36.8	797	21	AA609896	Human glycine tran
39	44.5	36.8	797	21	AA609897	Human glycine tran
40	44.5	36.8	797	21	AA623324	Human glycine tran
41	44.5	36.8	797	21	AA623327	Human glycine tran
42	44.5	36.8	797	21	AA623329	Human glycine tran
43	44	36.4	831	21	AA658002	Saccharomyces cere
44	44	36.4	273	23	ABP64378	Human ORF748. Hom
45	44	36.4	311	23	AA65734	Human G-protein co
	44	36.4	311	23	AA65735	Human G-protein co

ALIGNMENTS

RESULT 1
ID AAB68953
ID AAB68953 standard; peptide; 24 AA.
XX
AC AAB68953;
XX
DT 18-APR-2001 (first entry)
XX
DE Arabidopsis thaliana CDC27 protein conserved region #1.
XX
KW Cell cycle regulation; DNA replication; CDC7; CDC27A1; CDC27A2;
KW CDC27B; nematode resistance; endoreduplication; sterility;
KW polyploidy.
XX
OS Arabidopsis thaliana.
XX
FN WC200102430-A2.
XX
PD 11-JAN-2001.
XX
PF 05-JUL-2000; 2000WO-EP06401.
XX
PR 05-JUL-1999; 99BP-0202214.
XX
PA (CROP-) CROPDESIGN NV.
PA (UVRI-) UNIV RIO DE JANEIRO.
XX
FI Hemerly AS, Ferreira PCG, Rombauts S;
XX WPI; 2001-123101/13.
XX
PT Partially purified plant CDC27 or Cpc7 protein homolog, useful for

modulating DNA replication and for producing transgenic plants -
 Claim 1; Page 75; 86pp; English.

The present invention provides the protein and coding sequences of several Arabidopsis thaliana proteins which are involved in DNA replication and the regulation of the cell cycle. These include CDC7, CDC27A1, CDC27A2 and CDC27B. They are useful in the production of transgenic and mutant plants, as the mutations in the gene cause proteins to confer nematode resistance, sterility and polyploidy on plants and also lead to endoreduplication.

Note: This sequence is stated as being the same as that given as SEQ ID NO: 6 on page 6 of the specification. However, this sequence is shorter than that shown here.

Sequence 24 AA;
 Query Match 100.0%; Score 121; DB 22; Length 24;
 Best Local Similarity 100.0%; Pred. No. 6.2e-12;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNQLLARCYLNSQAYSAYILK 24
 |||||
 DB 1 VNQLLARCYLNSQAYSAYILK 24
 |||||

RESULT 2
 AAB68956
 ID AAB68956 standard; peptide; 716 AA.
 AC AAB68956;
 DT 18-APR-2001 (first entry)
 XX Arabidopsis thaliana CDC27A2 protein.
 DE Cell cycle regulation; DNA replication; CDC7; CDC27A1; CDC27A2;
 KW CDC27B; nematode resistance; endoreduplication; sterility;
 KW polyploidy.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 PN WO200102430-A2.
 PD 11-JAN-2001.
 PF 05-JUL-2000; 2000WO-EP06401.
 XX 05-JUL-1999; 99EP-0202214.
 PR 05-JUL-1999; 99EP-0202214.
 XX (CROP-) CROPDISEIGN NV.
 PA (UYRI-) UNIV RIO DE JANEIRO.
 PI Hemerly AS, Ferreira PCG, Rombauts S;
 DR WPI; 2001-123101/13.
 XX Partially purified plant CDC27 or CDC7 protein homolog, useful for
 PT modulating DNA replication and for producing transgenic plants -
 XX Claim 3; Page 78-81; 86pp; English.

The present invention provides the protein and coding sequences of several Arabidopsis thaliana proteins which are involved in DNA replication and the regulation of the cell cycle. These include CDC7, CDC27A1, CDC27A2 and CDC27B. They are useful in the production of transgenic and mutant plants, as the mutations in the gene cause proteins to confer nematode resistance, sterility and polyploidy on plants and also lead to endoreduplication.

Sequence 716 AA;
 Query Match 100.0%; Score 121; DB 22; Length 716;
 Best Local Similarity 100.0%; Pred. No. 6.2e-12;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 3e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNQLLARCYLNSQAYSAYILK 24
 |||||
 DB 37 VNQLLARCYLNSQAYSAYILK 60
 |||||

RESULT 3
 AAB68952
 ID AAB68952 standard; peptide; 728 AA.
 AC AAB68952;
 XX 18-APR-2001 (first entry)
 DT Arabidopsis thaliana CDC27A1 protein.
 DE Cell cycle regulation; DNA replication; CDC7; CDC27A1; CDC27A2;
 KW CDC27B; nematode resistance; endoreduplication; sterility;
 KW polyploidy.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 PN WO200102430-A2.
 PD 11-JAN-2001.
 PF 05-JUL-2000; 2000WO-EP06401.
 XX 05-JUL-1999; 99EP-0202214.
 PR (CROP-) CROPDISEIGN NV.
 PA (UYRI-) UNIV RIO DE JANEIRO.
 PI Hemerly AS, Ferreira PCG, Rombauts S;
 DR WPI; 2001-123101/13.
 XX Partially purified plant CDC27 or CDC7 protein homolog, useful for
 PT modulating DNA replication and for producing transgenic plants -
 XX Claim 3; Page 72-74; 86pp; English.

The present invention provides the protein and coding sequences of several Arabidopsis thaliana proteins which are involved in DNA replication and the regulation of the cell cycle. These include CDC7, CDC27A1, CDC27A2 and CDC27B. They are useful in the production of transgenic and mutant plants, as the mutations in the gene cause proteins to confer nematode resistance, sterility and polyploidy on plants and also lead to endoreduplication.

Sequence 728 AA;
 Query Match 100.0%; Score 121; DB 22; Length 728;
 Best Local Similarity 100.0%; Pred. No. 3e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VNQLLARCYLNSQAYSAYILK 24
 |||||
 DB 37 VNQLLARCYLNSQAYSAYILK 60
 |||||

RESULT 4
 AAB68960
 ID AAB68960 standard; peptide; 23 AA.
 AC AAB68960;
 XX 18-APR-2001 (first entry)
 DT Arabidopsis thaliana CDC27 protein conserved region #3.
 DE Arabidopsis thaliana CDC27 protein conserved region #3.
 XX

KW Cell cycle regulation; DNA replication; CDC7; CDC27A1; CDC27A2;
 KW CDC27B; nematode resistance; endoreduplication; sterility;
 KW polyploidy.

XX Arabidopsis thaliana.

XX WO200102430-A2.

XX 11-JAN-2001.

XX 05-JUL-2000; 2000WO-EP06401.

XX 05-JUL-1999; 99EP-0202214.

XX (CROP-) CROPDESIGN NV.

PA (UYRI-) UNIV RIO DE JANEIRO.

XX Hemerly AS, Ferreira PCG, Rombauts S;

XX WPI; 2001-123101/13.

XX Partially purified plant CDC27 or CDC7 protein homolog, useful for
 PT modulating DNA replication and for producing transgenic plants -

XX Claim 1; Page 6; 86pp; English.

XX The present invention provides the protein and coding sequences of
 CC several Arabidopsis thaliana proteins which are involved in DNA
 CC replication and the regulation of the cell cycle. These include CDC7,
 CC CDC27A1, CDC27A2 and CDC27B. They are useful in the production of
 CC transgenic and mutant plants, as the mutations in the gene cause
 CC proteins to confer nematode resistance, sterility and polyploidy on
 CC plants and also lead to endoreduplication.

CC Note: This sequence is stated as being the same as that given as SEQ ID
 CC NO: 6 in the sequence listing of the specification. However, this
 CC sequence is longer than that shown here.

XX Sequence 23 AA;

Query Match 88.0%; Score 106.5; DB 22; Length 23;

Best Local Similarity 95.8%; Pred. No. 1.1e-09;

Matches 23; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 VNQLLARCYSLSNQAYSAYYILK 24

DB 1 VNQLLARCYSLSN-QAYSAYYILK 23

RESULT 5

AAB68955

ID AAB68955 standard; peptide; 161 AA.

XX AAB68955;

XX 18-APR-2001 (first entry)

XX Arabidopsis thaliana CDC27B protein fragment.

XX Cell cycle regulation; DNA replication; CDC7; CDC27A1; CDC27A2;
 KW CDC27B; nematode resistance; endoreduplication; sterility;
 KW polyploidy.

XX Arabidopsis thaliana.

XX WO200102430-A2.

XX 11-JAN-2001.

XX 05-JUL-2000; 2000WO-EP06401.

XX 05-JUL-1999; 99EP-0202214.

XX (CROP-) CROPDESIGN NV.

PA (UYRI-) UNIV RIO DE JANEIRO.

XX Hemerly AS, Ferreira PCG, Rombauts S;

XX WPI; 2001-123101/13.

XX Partially purified plant CDC27 or CDC7 protein homolog, useful for
 PT modulating DNA replication and for producing transgenic plants -

XX Claim 1; Page 78; 86pp; English.

XX The present invention provides the protein and coding sequences of
 CC several Arabidopsis thaliana proteins which are involved in DNA
 CC replication and the regulation of the cell cycle. These include CDC7,
 CC CDC27A1, CDC27A2 and CDC27B. They are useful in the production of
 CC transgenic and mutant plants, as the mutations in the gene cause
 CC proteins to confer nematode resistance, sterility and polyploidy on
 CC plants and also lead to endoreduplication.

XX Sequence 161 AA;

Query Match 75.2%; Score 91; DB 22; Length 161;

Best Local Similarity 75.0%; Pred. No. 2.7e-06;

Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VNQLLARCYSLSNQAYSAYYILK 24

DB 36 VNQLLATSYLQNNQAYSAYHLLK 59

RESULT 6

ABG65520

ID ABG65520 standard; Protein; 162 AA.

XX ABG65520;

XX 23-AUG-2002 (first entry)

XX Cell cycle regulation protein HOBBIT (HBT) #1.

XX Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
 KW pattern formation; Plant meristem development; plant yield; drought;
 KW seedling emergency; root generation; shade avoidance response;
 KW vascular strand formation; patterning; parthenocarpic fruit;
 KW cell cycle regulation; endoreduplication; cell division;
 KW transgenic plant; plant development; plant morphology; plant physiology;
 KW plant biochemistry.

XX Arabidopsis thaliana.

XX WO200238599-A2.

XX 16-MAY-2002.

XX 13-NOV-2001; 2001WO-EPI3116.

XX 13-NOV-2000; 2000EP-0870271.

XX 30-NOV-2000; 2000US-250402P.

XX (UYUT-) RIJKSUNIV UTRECHT.

XX Scheres BJG, Bilou I, Folmer SDH;

XX WPI; 2002-490065/52.

XX N-PSDB; ABK93343.

XX Use of plant cdc27B for modulating or mimicking auxin-related effects
 PT in plants or plant cells, or for regulation of cell cycle of plant cell
 PT -

XX Claim 30; Page 144-145; 207pp; English.

XX The invention described the use of a plant cdc27B (also termed HOBBIT

CC (HBT)) useful for modulating or mimicking auxin-related effects in a
CC plant or plant cell. The modulation or mimicking of auxin-related effects
CC results in: altered cell fate and/or altered pattern formation in a plant
CC or plant cell; alteration in the size and/or number of naturally
CC occurring plant meristems; modification of the rate of organ or tissue emanation
CC tissues, and/or a modification of the rate of organ or tissue emanation
CC from a plant meristem, and/or a modification of the arrangement of organs
CC and/or tissues in a plant; increased survival rate under drought conditions;
CC rate of plants; an enhanced survival rate under drought conditions;
CC increase of seedling emergency; mimicking of root generation in tissue
CC cultures; an increased shade avoidance response; altering vascular strand
CC formation and patterning in a plant; or production of parthenocarpic
CC fruits. HBT also regulates the cell cycle of the plant cell results in:
CC modulated endoreduplication in a plant; modulation of sterility in
CC plants; increasing the cell cycle or increase of the rate of cell
CC division; alteration in a number of organs or tissues, and/or the modification of
CC the rate of organ or tissue emanation from a plant meristem and/or
CC modification of the arrangement of the organs and/or tissues in a plant
CC and/or an increased plant yield or enhancement of survival rate of
CC plants. A nucleic acid encoding HBT is useful for producing transgenic
CC plants, plant cells or plant tissues and in the regeneration of a plant
CC from a plant cell. HBT nucleic acids and proteins are useful for
CC modifying cell fate, pattern formation, plant development, plant
CC morphology, plant physiology and/or plant biochemistry. This is the amino
CC acid sequence of an Arabidopsis thaliana HOBBIT (HBT) or cdc27B protein
CC described in the invention.
XX SQ Sequence 162 AA;

Query Match 75.2%; Score 91; DB 23; Length 162;
Best Local Similarity 75.0%; Pred. No. 2.7e-06;
Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 VNQLLARCYSNQAYSAYILK 24
||| ||| ||| : ||| ||| : |||
DB 36 VNQLLATSYLQNNQAYSAYHLK 59

RESULT 7
ABG65521
ID ABG65521 standard; Protein; 166 AA.
XX AC ABG65521;
XX DT 23-AUG-2002 (first entry)
XX DE Cell cycle regulation protein HOBBIT (HBT) #2.
XX KW plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
KW pattern formation; plant meristem development; plant yield; drought;
KW seedling emergency; root generation; shade avoidance response;
KW vascular strand formation; patterning; parthenocarpic fruit;
KW cell cycle regulation; endoreduplication; cell division;
KW transgenic plant; plant development; plant morphology; plant physiology;
KW plant biochemistry.
XX OS Arabidopsis thaliana.
XX PN WO200238599-A2.
XX PD 16-MAY-2002.
XX PF 13-NOV-2001; 2001WO-EPI3116.
XX PR 13-NOV-2000; 2000EP-0870271.
XX PR 30-NOV-2000; 2000US-250402P.
XX PA (UYUT-) RIJKSUNIV UTRECHT.
XX PI Scheres BJG, Billaou I, Folmer SDH;
XX WPI; 2002-490065/52.

DR N-PSDB; ABK93344.
XX Use of plant cdc27B for modulating or mimicking auxin-related effects
PT in plants or plant cells, or for regulation of cell cycle of plant cell
PT -
XX Claim 30; Page 145; 207pp; English.
XX The invention described the use of a plant cdc27B (also termed HOBBIT
CC (HBT)) useful for modulating or mimicking auxin-related effects in a
CC plant or plant cell. The modulation or mimicking of auxin-related effects
CC results in: altered cell fate and/or altered pattern formation in a plant
CC or plant cell; alteration in the size and/or number of naturally
CC occurring plant meristems; modification of the numbers of organs or
CC tissues, and/or a modification of the rate of organ or tissue emanation
CC from a plant meristem, and/or a modification of the arrangement of organs
CC and/or tissues in a plant; increased survival rate under drought conditions;
CC rate of plants; an enhanced survival rate under drought conditions;
CC increase of seedling emergency; mimicking of root generation in tissue
CC cultures; an increased shade avoidance response; altering vascular strand
CC formation and patterning in a plant; or production of parthenocarpic
CC fruits. HBT also regulates the cell cycle of the plant cell results in:
CC modulated endoreduplication in a plant; modulation of sterility in
CC plants; increasing the cell cycle or increase of the rate of cell
CC division; alteration in a number of organs or tissues, and/or the modification of
CC the rate of organ or tissue emanation from a plant meristem and/or
CC modification of the arrangement of the organs and/or tissues in a plant
CC and/or an increased plant yield or enhancement of survival rate of
CC plants. A nucleic acid encoding HBT is useful for producing transgenic
CC plants, plant cells or plant tissues and in the regeneration of a plant
CC from a plant cell. HBT nucleic acids and proteins are useful for
CC modifying cell fate, pattern formation, plant development, plant
CC morphology, plant physiology and/or plant biochemistry. This is the amino
CC acid sequence of an Arabidopsis thaliana HOBBIT (HBT) or cdc27B protein
CC described in the invention.
XX SQ Sequence 166 AA;

Query Match 75.2%; Score 91; DB 23; Length 166;
Best Local Similarity 75.0%; Pred. No. 2.8e-06;
Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 VNQLLARCYSNQAYSAYILK 24
||| ||| ||| : ||| ||| : |||
DB 36 VNQLLATSYLQNNQAYSAYHLK 59

RESULT 8
ABG65529
ID ABG65529 standard; Protein; 518 AA.
XX AC ABG65529;
XX DT 23-AUG-2002 (first entry)
XX DE Functionally inactive HOBBIT (HBT) protein #6.
XX KW plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
KW pattern formation; plant meristem development; plant yield; drought;
KW seedling emergency; root generation; shade avoidance response;
KW vascular strand formation; patterning; parthenocarpic fruit;
KW cell cycle regulation; endoreduplication; cell division;
KW transgenic plant; plant development; plant morphology; plant physiology;
KW plant biochemistry.
XX OS Arabidopsis thaliana.
XX PN WO200238599-A2.
XX PD 16-MAY-2002.
XX PF 13-NOV-2001; 2001WO-EPI3116.

AA668958	ID	AAB68958 standard; peptide; 739 AA.
XX	AC	AAB68958;
XX	DT	18-APR-2001 (first entry)
XX	DE	Arabidopsis thaliana CDC27B protein.
XX	KW	Cell cycle regulation; DNA replication; CDC7; CDC27A1; CDC27A2;
KW	KW	CDC27B; nematode resistance; endoreduplication; sterility;
KW	KW	polyploidy.
XX	OS	Arabidopsis thaliana.
XX	PN	WO200102430-A2.
XX	PD	11-JAN-2001.
XX	PF	05-JUL-2000; 2000WO-EP06401.
XX	PR	05-JUL-1999; 99EP-0202214.
XX	PA	(CROP-) CROPDESIGN NV.
PA	PA	(UYRI-) UNIV RIO DE JANEIRO.
XX	PI	Hemerly AS, Ferreira PCG, Rombauts S;
XX	DR	WPI; 2001-123101/13.
XX	PT	Partially purified plant CDC27 or CDC7 protein homolog, useful for
PT	PT	modulating DNA replication and for producing transgenic plants -
XX	PS	Claim 3; Page 82-84; 86pp; English.
XX	CC	The present invention provides the protein and coding sequences of
CC	CC	several Arabidopsis thaliana proteins which are involved in DNA
CC	CC	replication and the regulation of the cell cycle. These include CDC7,
CC	CC	CDC27A1, CDC27A2 and CDC27B. They are useful in the production of
CC	CC	transgenic and mutant plants, as the mutations in the gene cause
CC	CC	proteins to confer nematode resistance, sterility and polyploidy on
CC	CC	plants and also lead to endoreduplication.
XX	SQ	Sequence 739 AA;
Query Match 75.2%; Score 91; DB 22; Length 739;		
Best Local Similarity 75.0%; Pred. No. 1.5e-05;		
Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps		
Qy	1	VNIQLLARCYLNSQAYSAYILK 24
Db	36	VNIQLLATSVLQNNQAYSAYHLK 59
RESULT 12		
ABG65522	ID	ABG65522 standard; Protein; 740 AA.
XX	AC	ABG65522;
XX	DT	23-AUG-2002 (first entry)
XX	DE	Cell cycle regulation protein HOBBIT (HBT) #3.
KW	KW	Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
KW	KW	pattern formation; plant meristem development; plant yield; drought;
KW	KW	seedling emergency; root generation; shade avoidance response;
KW	KW	vascular strand formation; patterning; parthenocarpic fruit;
KW	KW	cell cycle regulation; endoreduplication; cell division;
KW	KW	transgenic plant; plant development; plant morphology; plant physiology;
KW	KW	plant biochemistry.
XX	OS	Arabidopsis thaliana.

XX KW WO200238599-A2.
 PN
 XX
 XX
 PD 16-MAY-2002.
 XX
 XX 13-NOV-2001; 2001WO-EPI3116.
 XX
 XX 13-NOV-2000; 2000EP-0870271.
 PR 30-NOV-2000; 2000US-250402P.
 XX
 XX (UYUT-) RIJKSUNIV UTRECHT.
 PA
 XX Scheres BJG, Blilou I, Folmer SDH;
 XX WPI; 2002-490065/52.
 XX
 XX Use of plant cdc27B for modulating or mimicking auxin-related effects
 PT in plants or plant cells, or for regulation of cell cycle of plant cell
 PT
 PS Claim 30; Page 146-149; 207pp; English.
 XX
 XX The invention described the use of a plant cdc27B (also termed HOBBIT
 CC (HBT)) useful for modulating or mimicking auxin-related effects in a
 CC plant or plant cell. The modulation or mimicking of auxin-related effects
 CC results in: altered cell fate and/or altered pattern formation in a plant
 CC or plant cell; alteration in the size and/or number of naturally
 CC occurring plant meristems; modification of the numbers of organs or
 CC tissues, and/or a modification of the rate of organ or tissue emanation
 CC from a plant meristem, and/or a modification of the arrangement of organs
 CC and/or tissues in a plant; increased plant yield; an increased survival
 CC rate of plants; an enhanced survival rate under drought conditions;
 CC increase of seedling emergency; mimicking of root generation in tissue
 CC cultures; an increased shade avoidance response; altering vascular strand
 CC formation and patterning in a plant; or production of parthenocarpic
 CC fruits. HBT also regulates the cell cycle of the plant cell results in:
 CC modulated endoreduplication in a plant; modulation of sterility in
 CC plants; increasing the cell cycle or increase of the rate of cell
 CC division; alteration in the size of naturally occurring meristems; a
 CC modification in a number of organs or tissues, and/or the modification of
 CC the rate of organ or tissue emanation from a plant meristem and/or
 CC modification of the arrangement of the organs and/or tissues in a plant
 CC and/or an increased plant yield or enhancement of survival rate of
 CC plants. A nucleic acid encoding HBT is useful for producing transgenic
 CC plants, plant cells or plant tissues and in the regeneration of a plant
 CC from a plant cell. HBT nucleic acids and proteins are useful for
 CC modifying cell fate, pattern formation, plant development, plant
 CC morphology, plant physiology and/or plant biochemistry. This is the amino
 CC acid sequence of an Arabidopsis thaliana HOBBIT (HBT) or cdc27B protein
 CC described in the invention.
 XX
 XX Sequence 740 AA;
 XX
 XX Query Match 75.2%; Score 91; DB 23; Length 740;
 XX Best Local Similarity 75.0%; Pred. No. 1.6e-05;
 XX Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VNQLLARCYLNSQAYSAYILK 24
 DB 36 VNQLLATSYLQNNQAYSAYHLK 59
 RESULT 13
 ABG65524
 ID ABG65524 standard; Protein; 743 AA.
 XX
 XX AC ABG65524;
 XX
 XX 23-AUG-2002 (first entry)
 DT
 XX Functionally inactive HOBBIT (HBT) protein #1.
 DE
 XX Plant; cdc27B; cyclin; HOBBIT, HBT; auxin-related effect; cell fate;

KW pattern formation; plant meristem development; plant yield; drought;
 KW seedling emergency; root generation; shade avoidance response;
 KW vascular strand formation; patterning; parthenocarpic fruit;
 KW cell cycle regulation; endoreduplication; cell division;
 KW transgenic plant; plant development; plant morphology; plant physiology;
 KW plant biochemistry.
 XX Arabidopsis thaliana.
 OS WO200238599-A2.
 PN 16-MAY-2002.
 PD
 XX 13-NOV-2001; 2001WO-EPI3116.
 PF 13-NOV-2000; 2000EP-0870271.
 PR 30-NOV-2000; 2000US-250402P.
 XX
 XX (UYUT-) RIJKSUNIV UTRECHT.
 PA
 XX Scheres BJG, Blilou I, Folmer SDH;
 PI WPI; 2002-490065/52.
 XX
 XX Use of plant cdc27B for modulating or mimicking auxin-related effects
 PT in plants or plant cells, or for regulation of cell cycle of plant cell
 PT
 PS Claim 36; Page 178-181; 207pp; English.
 XX
 XX The invention described the use of a plant cdc27B (also termed HOBBIT
 CC (HBT)) useful for modulating or mimicking auxin-related effects in a
 CC plant or plant cell. The modulation or mimicking of auxin-related effects
 CC results in: altered cell fate and/or altered pattern formation in a plant
 CC or plant cell; alteration in the size and/or number of naturally
 CC occurring plant meristems; modification of the numbers of organs or
 CC tissues, and/or a modification of the rate of organ or tissue emanation
 CC from a plant meristem, and/or a modification of the arrangement of organs
 CC and/or tissues in a plant; increased plant yield; an increased survival
 CC rate of plants; an enhanced survival rate under drought conditions;
 CC increase of seedling emergency; mimicking of root generation in tissue
 CC cultures; an increased shade avoidance response; altering vascular strand
 CC formation and patterning in a plant; or production of parthenocarpic
 CC fruits. HBT also regulates the cell cycle of the plant cell results in:
 CC modulated endoreduplication in a plant; modulation of sterility in
 CC plants; increasing the cell cycle or increase of the rate of cell
 CC division; alteration in the size of naturally occurring meristems; a
 CC modification in a number of organs or tissues, and/or the modification of
 CC the rate of organ or tissue emanation from a plant meristem and/or
 CC modification of the arrangement of the organs and/or tissues in a plant
 CC and/or an increased plant yield or enhancement of survival rate of
 CC plants. A nucleic acid encoding HBT is useful for producing transgenic
 CC plants, plant cells or plant tissues and in the regeneration of a plant
 CC from a plant cell. HBT nucleic acids and proteins are useful for
 CC modifying cell fate, pattern formation, plant development, plant
 CC morphology, plant physiology and/or plant biochemistry. This is the amino
 CC acid sequence of a functionally inactive Arabidopsis thaliana HOBBIT
 CC (HBT) or cdc27B protein described in the invention.
 XX
 XX Sequence 743 AA;
 XX
 XX Query Match 75.2%; Score 91; DB 23; Length 743;
 XX Best Local Similarity 75.0%; Pred. No. 1.6e-05;
 XX Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VNQLLARCYLNSQAYSAYILK 24
 DB 36 VNQLLATSYLQNNQAYSAYHLK 59
 RESULT 14
 ABG65523
 ID ABG65523 standard; Protein; 744 AA.

XX AC ABG65523;
 XX DT 23-AUG-2002 (first entry)
 XX DE Cell cycle regulation protein HOBBIT (HBT) #4.
 XX DE Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
 KW pattern formation; plant meristem development; plant yield; drought;
 KW seedling emergency; root generation; shade avoidance response;
 KW vascular strand formation; patterning; parthenocarpic fruit;
 KW cell cycle regulation; endoreduplication; cell division;
 KW transgenic plant; plant development; plant morphology; plant physiology;
 KW plant biochemistry.
 XX OS Arabidopsis thaliana.
 XX PN WO200238599-A2.
 XX PD 16-MAY-2002.
 XX PF 13-NOV-2001; 2001WO-EPI3116.
 XX PR 13-NOV-2000; 2000EP-0870271.
 XX PR 30-NOV-2000; 2000US-250402P.
 XX PA (UYUT-) RIJKSUNIV UTRECHT.
 XX PI Scheres BJG, Blilou I, Folmer SDH;
 XX DR WPI; 2002-490065/52.
 XX PT Use of plant cdc27B for modulating or mimicking auxin-related effects
 XX PT in plants or plant cells, or for regulation of cell cycle of plant cell
 XX PS Claim 30; Page 149-152; 207pp; English.
 XX CC The invention described the use of a plant cdc27B (also termed HOBBIT
 CC (HBT)) useful for modulating or mimicking auxin-related effects in a
 CC plant or plant cell. The modulation or mimicking of auxin-related effects
 CC results in: altered cell fate and/or altered pattern formation in a plant
 CC or plant cell; alteration in the size and/or number of naturally
 CC occurring plant meristems; modification of the numbers of organs or
 CC tissues, and/or a modification of the rate of organ or tissue emanation
 CC from a plant meristem, and/or a modification of the arrangement of organs
 CC and/or tissues in a plant; increased plant yield; an increased survival
 CC rate of plants; an enhanced emergency; mimicking of root generation in tissue
 CC cultures; an increased shade avoidance response; altering vascular strand
 CC formation and patterning in a plant; or production of parthenocarpic
 CC fruits. HBT also regulates the cell cycle of the plant cell results in:
 CC modulated endoreduplication in a plant; modulation of sterility in
 CC plants; increasing the cell cycle or increase of the rate of cell
 CC division; alteration in the size of naturally occurring meristems; a
 CC modification in a number of organs or tissues, and/or the modification of
 CC the rate of organ or tissue emanation from a plant meristem and/or
 CC modification of the arrangement of the organs and/or tissues in a plant
 CC and/or an increased plant yield or enhancement of survival rate of
 CC plants. A nucleic acid encoding HBT is useful for producing transgenic
 CC plants, plant cells or plant tissues and in the regeneration of a plant
 CC from a plant cell. HBT nucleic acids and proteins are useful for
 CC modifying cell fate, pattern formation, plant development, plant
 CC morphology, plant physiology and/or plant biochemistry. This is the amino
 CC acid sequence of an Arabidopsis thaliana HOBBIT (HBT) or cdc27B protein
 CC described in the invention.
 XX SQ Sequence 744 AA;
 Query Match 75.2%; Score 91; DB 23; Length 744;
 Best Local Similarity 75.0%; Pred. No. 1.6e-05;
 Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VNQLLARCYLNSQAYSAYILK 24
 DB 36 VNQLLATSYLQNNQAYSAYHLK 59
 RESULT 15
 ABG65526
 ID ABG65526 standard; Protein; 744 AA.
 XX AC ABG65526;
 XX DT 23-AUG-2002 (first entry)
 XX DE Functionally inactive HOBBIT (HBT) protein #3.
 XX DE Plant; cdc27B; cyclin; HOBBIT; HBT; auxin-related effect; cell fate;
 KW pattern formation; plant meristem development; plant yield; drought;
 KW seedling emergency; root generation; shade avoidance response;
 KW vascular strand formation; patterning; parthenocarpic fruit;
 KW cell cycle regulation; endoreduplication; cell division;
 KW transgenic plant; plant development; plant morphology; plant physiology;
 KW plant biochemistry.
 XX OS Arabidopsis thaliana.
 XX PN WO200238599-A2.
 XX PD 16-MAY-2002.
 XX PF 13-NOV-2001; 2001WO-EPI3116.
 XX PR 13-NOV-2000; 2000EP-0870271.
 XX PR 30-NOV-2000; 2000US-250402P.
 XX PA (UYUT-) RIJKSUNIV UTRECHT.
 XX PI Scheres BJG, Blilou I, Folmer SDH;
 XX DR WPI; 2002-490065/52.
 XX PT Use of plant cdc27B for modulating or mimicking auxin-related effects
 XX PT in plants or plant cells, or for regulation of cell cycle of plant cell
 XX PS Claim 36; Page 185-188; 207pp; English.
 XX CC The invention described the use of a plant cdc27B (also termed HOBBIT
 CC (HBT)) useful for modulating or mimicking auxin-related effects in a
 CC plant or plant cell. The modulation or mimicking of auxin-related effects
 CC results in: altered cell fate and/or altered pattern formation in a plant
 CC or plant cell; alteration in the size and/or number of naturally
 CC occurring plant meristems; modification of the numbers of organs or
 CC tissues, and/or a modification of the rate of organ or tissue emanation
 CC from a plant meristem, and/or a modification of the arrangement of organs
 CC and/or tissues in a plant; increased plant yield; an increased survival
 CC rate of plants; an enhanced emergency; mimicking of root generation in tissue
 CC cultures; an increased shade avoidance response; altering vascular strand
 CC formation and patterning in a plant; or production of parthenocarpic
 CC fruits. HBT also regulates the cell cycle of the plant cell results in:
 CC modulated endoreduplication in a plant; modulation of sterility in
 CC plants; increasing the cell cycle or increase of the rate of cell
 CC division; alteration in the size of naturally occurring meristems; a
 CC modification in a number of organs or tissues, and/or the modification of
 CC the rate of organ or tissue emanation from a plant meristem and/or
 CC modification of the arrangement of the organs and/or tissues in a plant
 CC and/or an increased plant yield or enhancement of survival rate of
 CC plants. A nucleic acid encoding HBT is useful for producing transgenic
 CC plants, plant cells or plant tissues and in the regeneration of a plant
 CC from a plant cell. HBT nucleic acids and proteins are useful for
 CC modifying cell fate, pattern formation, plant development, plant
 CC morphology, plant physiology and/or plant biochemistry. This is the amino
 CC acid sequence of a functionally inactive Arabidopsis thaliana HOBBIT

CC (HBT) or cdc27B protein described in the invention.

XX

SQ Sequence 744 AA;

Query Match 75.2%; Score 91; DB 23; Length 744;

Best Local Similarity 75.0%; Pred. No. 1.6e-05;

Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 VNLQLLARCYSLSQAYSAYYILK 24

Db 36 VNLQLLATSYLQNNQAYSAYHLK 59

Search completed: February 2, 2004, 13:01:20

Job time : 649 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 2, 2004, 13:25:05 ; Search time 4814 Seconds
(without alignments)
121.169 Million cell updates/sec

Title: US-10-036-492-6
Perfect score: 121
Sequence: 1 VNQLLARCYSLSQAYSYILK 24

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2/1/USPTO_spool/US10036492/runat_02022004_085034_13467/app_query.fasta_1.199
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10036492@cgn 1 1 2810 @runat_02022004_085034_13467 -NCFU=6 -ICPU=3
-NO.MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic1:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	116	95.9	554	28	BH593533 BOGEA03TF
C 2	116	95.9	619	28	BH428018 BOHLN95TR
C 3	116	95.9	767	28	BH603761 BOGZJ78TF
C 4	99	81.8	390	13	BQ140592 NFQ037H06P
5	99	81.8	527	13	BH93652 P080F01 P
6	99	81.8	552	9	AW267890 EST306168
7	99	81.8	656	13	BQ138306 NFQ01G11P
8	91	75.2	622	28	B78168 T31120TF TA
9	89	73.6	467	12	B1176998 EST517943
10	89	73.6	710	10	BG593671 EST492349
11	89	73.6	710	12	B1178043 EST519888
12	89	73.6	727	12	BG887406 EST513257
13	89	73.6	742	12	BQ047257 EST596375
C 14	85	70.2	639	29	AG226554 Lotus_jap
15	63	52.1	402	14	CB366511
16	63	52.1	497	14	CB363774
17	63	52.1	593	10	BG737576
18	63	52.1	639	13	BQ258401 fz79c02.Y
19	63	52.1	641	13	BQ262043 fz70b08.Y
20	62	51.2	378	13	BY029612
C 21	61	50.4	705	9	AW342355 GthST212
C 22	61	50.4	735	9	AW342276 GthST28
C 23	61	50.4	932	9	AW342275 GthST27
24	60	49.6	362	13	BY193662
25	60	49.6	597	14	CB583162 AMGNNUC:C
26	60	49.6	756	10	BZ279064 G01156255
27	59	48.8	321	13	BY331052
28	59	48.8	328	13	BY144969
29	59	48.8	345	13	BY325168
30	59	48.8	348	14	CA407596
31	59	48.8	356	13	BY013739
32	59	48.8	360	13	BY324782
33	59	48.8	364	13	BY146525
34	59	48.8	366	14	CA875900
35	59	48.8	381	13	BY015764
36	59	48.8	384	10	B8844029
37	59	48.8	397	13	BY075754
38	59	48.8	406	14	M78440
39	59	48.8	414	10	BZ226683
40	59	48.8	418	14	T82077
41	59	48.8	443	14	CA543709
42	59	48.8	518	10	BG503879
43	59	48.8	528	14	CA886679
44	59	48.8	547	4	BX520775
45	59	48.8	549	12	BJ095805

ALIGNMENTS

RESULT 1
BH593533/c
LOCUS
DEFINITION BOGEA03TF BOGE Brassica oleracea genomic clone BOGEA03, genomic
survey sequence.
ACCESSION BH593533
VERSION BH593533.1 GI:17845985
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.

BH593533 554 bp DNA linear GSS 15-DEC-2001
BOGEA03TF BOGE Brassica oleracea genomic clone BOGEA03, genomic
survey sequence.

BH593533
BH593533.1 GI:17845985
GSS.
Brassica oleracea
Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 554)
 AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished
 COMMENT Other GSSs: BOGEA03TR
 Contact: Chris Town

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TF
 Class: sheared ends.

FEATURES

source

Location/Qualifiers
 1..554
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BOGEA03"
 /clone_lib="BOGE"
 /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT 204 a 99 c 138 g 113 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.67e-07 Length: 554
 Score: 116.00 Matches: 23
 Percent Similarity: 95.83% Conservative: 0
 Best Local Similarity: 95.83% Mismatches: 1
 Query Match: 95.87% Indels: 0
 DB: 28 Gaps: 0

US-10-036-492-6 (1-24) x BH593533 (1-554)

QY 1 ValAsnLeuGlnLeuAlaAAGCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 Db 102 GTGAACCTGCAATTTAGTCTAGTGTACTTAAGTAACAGTCAACCTTATAGTCTTAT 43
 QY 21 Tyrlleleulys 24
 Db 42 TACATCCTCAAA 31

RESULT 2
 BH428018/c 619 bp DNA linear GSS 12-DEC-2001
 LOCUS BOHLN95TR BOHL Brassica oleracea genomic clone BOHLN95, genomic
 DEFINITION survey sequence.

ACCESSION BH428018
 VERSION BH428018.1 GI:17613746
 KEYWORDS GSS.
 SOURCE Brassica oleracea

ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 619)
 AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished
 COMMENT Other GSSs: BOHLN95TF
 Contact: Chris Town

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TF
 Class: sheared ends.

FEATURES

Location/Qualifiers

source

1..619
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BOHLN95"
 /clone_lib="BOHL"
 /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT 245 a 102 c 124 g 148 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.95e-07 Length: 619
 Score: 116.00 Matches: 23
 Percent Similarity: 95.83% Conservative: 0
 Best Local Similarity: 95.83% Mismatches: 1
 Query Match: 95.87% Indels: 0
 DB: 28 Gaps: 0

US-10-036-492-6 (1-24) x BH428018 (1-619)

QY 1 ValAsnLeuGlnLeuAlaAAGCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 Db 422 GTGAACCTGCAATTTAGTCTAGTGTACTTAAGTAACAGTCAACCTTATAGTCTTAT 363

QY 21 Tyrlleleulys 24

Db 362 TACATCCTCAAA 351

RESULT 3

BH603761/c

LOCUS BOGZJ78TF BOGZ Brassica oleracea genomic clone BOGZJ78, genomic

DEFINITION survey sequence.

ACCESSION BH603761

VERSION BH603761.1 GI:17856207

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 767)

AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.

TITLE Whole genome shotgun sequencing of Brassica oleracea

JOURNAL Unpublished

COMMENT Other GSSs: BOGZJ78TR

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..767

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TO1000DH3"

/db_xref="taxon:3712"

/clone="BOGZJ78"

/clone_lib="BOGZ"

/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared

genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT 294 a 131 c 162 g 180 t

ORIGIN

Alignment Scores:

Pred. No.: 2.64e-07 Length: 767

Score: 116.00 Matches: 23

Percent Similarity: 95.83% Conservative: 0
 Best Local Similarity: 95.83% Mismatches: 1
 Query Match: 95.87% Indels: 0
 DB: 28 Gaps: 0

US-10-036-492-6 (1-24) x BH603761 (1-767)

QY 1 ValAsnLeuGlnLeuLeuAlaAArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
 Db 440 GTGAACCTGCAATGTTAGTGTAGTGTACTTAAAGTCAACGTCACCTTATAGTCTTAT 381

QY 21 TyrIleLeuLys 24
 Db 380 TACATCCTCAAA 369

RESULT 4
 BQ140592 390 bp mRNA linear EST 26-APR-2002
 LOCUS NF037H06PH1F060 Phoma-infected Medicago truncatula cDNA clone
 DEFINITION NF037H06PH 5', mRNA sequence.
 ACCESSION BQ140592
 VERSION BQ140592.1 GI:20276718
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 REFERENCE 1 (bases 1 to 390)
 AUTHORS Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,
 Gonzales,R.A., Bell,C.J., Inman,J.T., Waugh,M.E., Sullivan,J.P.,
 May,G.D. and Paiva,N.L.
 TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula Phoma-infected library
 JOURNAL Unpublished
 COMMENT Contact: Paiva NL
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7317
 Fax: 580 221 7380
 Email: nlpaiva@cnhble.org
 Insert Length: 390 Std Error: 0.00
 Plate: 037 Row: H Column: 06
 Seq primer: TCACACAGGAACAGCTATGAC.
 FEATURES
 source
 1..390
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /db_xref="taxon:3880"
 /clone="NF037H06PH"
 /tissue_type="leaf"
 /dev_stage="Pathogen-induced, young trifoliolate"
 /clone_lib="Phoma-infected"
 /note="Vector: pBluescript SK(-); Young trifoliolate leaves
 of Medicago truncatula were excised and dip-inoculated in
 a spore suspension of Phoma medicaginis, and incubated in
 humid dishes. Pools of leaves were harvested at 0, 15, and
 30 minutes and 1, 2, 3, 6, 14, 24, 48, 72, and 96 hours,
 and used to prepare total RNA. cDNA was prepared from
 polyA+ enriched, pooled samples of equivalent amounts of
 total RNA from each sample. The cDNA was directionally
 ligated into the Uni-Zap XR vector (Stratagene) and
 packaged using the Gigapack III Gold packaging extracts.
 Phagemids containing cDNA inserts were in vivo excised
 from the recombinant Uni-Zap XR vector using ExAest
 helper phage and the E. coli strain XL1-Blue MRF'
 (Stratagene). Excised plasmids were plated using SOLR
 cells."

117 a 90 c 70 g 113 t

BASE COUNT
 ORIGIN

Alignment Scores:
 Pred. No.: 5.65e-05 Length: 527
 Score: 99.00 Matches: 19
 Percent Similarity: 91.30% Conservative: 2
 Best Local Similarity: 82.61% Mismatches: 2
 Query Match: 81.82% Indels: 0
 DB: 13 Gaps: 0

US-10-036-492-6 (1-24) x BQ140592 (1-390)

QY 2 AsnLeuGlnLeuLeuAlaAArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 21
 Db 131 AATTGCAATTATTAGCTGCTGTTACTTGCAGATAATCAAGCTTATTCGCATACAT 190

QY 22 IleLeuLys 24
 Db 191 ATTCTAAAG 199

RESULT 5
 BQ93652 527 bp mRNA linear EST 17-OCT-2002
 LOCUS BU93652 Populus petioles cDNA library Populus tremula cDNA 5 prime,
 DEFINITION mRNA sequence.
 ACCESSION BU93652
 VERSION BU93652.1 GI:24104717
 KEYWORDS EST.
 SOURCE Populus tremula
 ORGANISM Populus tremula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Malpighiales; Salicaceae; Populus.
 REFERENCE 1 (bases 1 to 527)
 AUTHORS Umeberg,P., Bhalarao,R.R., Jansson,S. and Sterky,F.
 TITLE The Poplar tree transcriptome: Analysis of expressed sequence tags
 from multiple libraries
 JOURNAL Unpublished
 COMMENT Contact: BHALERAO RUPALI R.
 Umea Plant Science Center
 Department of Plant Physiology
 University of Umea, 901 87 Umea, Sweden
 Tel: +46 90 786 5279
 Fax: +46 90 786 6676
 Email: rupali.bhalarao@plantphys.umu.se.
 FEATURES
 source
 1..527
 /organism="Populus tremula"
 /mol_type="mRNA"
 /db_xref="taxon:113636"
 /tissue_type="petioles"
 /clone_lib="Populus petioles cDNA library"

BASE COUNT 150 a 132 c 88 g 157 t
 ORIGIN

Alignment Scores:
 Pred. No.: 5.65e-05 Length: 527
 Score: 99.00 Matches: 19
 Percent Similarity: 91.30% Conservative: 2
 Best Local Similarity: 82.61% Mismatches: 2
 Query Match: 81.82% Indels: 0
 DB: 13 Gaps: 0

US-10-036-492-6 (1-24) x BU93652 (1-527)

QY 2 AsnLeuGlnLeuLeuAlaAArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 21
 Db 362 AATTGCAATTCTAGCAGCGTGTATCTGCAGAAAATAATCAAGCATCTCTCGCATCAT 421

QY 22 IleLeuLys 24
 Db 422 ATCTAAAA 430

RESULT 6
 AW267890

```

LOCUS      AW267890                      552 bp    mRNA    linear    EST 07-SEP-2000
DEFINITION EST306168 DSIR Medicago truncatula cDNA clone pDSIR-802, mRNA
ACCESSION  AW267890
VERSION    AW267890.1 GI:5654846
KEYWORDS   EST.
SOURCE     Medicago truncatula (barrel medic)
ORGANISM   Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE  1 (bases 1 to 552)
AUTHORS   Pedorova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng
            ,H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
            Holt,I.E. and Fraser,C.M.
TITLE     ESTs from roots of Medicago truncatula after inoculation with
            Phytophthora medicaginis
JOURNAL    Unpublished
COMMENT    Contact: Carroll P. Vance
            Department of Agronomy and Plant Genetics
            University of Minnesota
            411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA
            Tel: 612 625 5715
            Fax: 651-649-5058
            Email: vance004@maroon.tc.umn.edu
            TIGR sequence name: M250399e
            More information, including clone ordering, is available at:
            http://chryslie.tamu.edu/medicago
            Seq primer: SKmod (CfA gMA CTA 5'g GAT CC).
FEATURES   source
            1..552
                Location/Qualifiers
                /organism="Medicago truncatula"
                /mol_type="mRNA"
                /cultivar="genotype Al7"
                /db_xref="taxon:3880"
                /clone="pDSIR-802"
                /tissue_type="roots infected with Phytophthora
                medicaginis"
                /dev_stage="roots harvested at 10 days post inoculation
                with Phytophthora medicaginis"
                /lab_host="E. coli strain XL0LR"
                /clone_lib="DSIR"
                /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
                XhoI; cDNA was prepared from polyA+ enriched RNA from
                roots harvested at 10 days post inoculation with
                Phytophthora medicaginis. The cDNA was directionally
                ligated into the Uni-ZAP XR vector from Stratagene and
                packaged using Gigapack III Gold packaging extracts.
                Plasmids containing cDNA inserts were excised from the
                recombinant lambda-ZAP phage using Ex-Assist helper phage
                and propagated in XL0LR cells. Note: EST may be of fungal
                origin."
BASE COUNT  158 a 111 c 119 g 164 t
ORIGIN
Alignment Scores:
Pred. No.: 6.03e-05 Length: 552
Score: 99.00 Matches: 19
Percent Similarity: 91.30% Conservatve: 2
Best Local Similarity: 82.61% Mismatches: 2
Query Match: 81.82% Indels: 0
DB: 9 Gaps: 0

US-10-036-492-6 (1-24) x AW267890 (1-552)

Qy 2 AsnLeuGlnLeuLeuAlaAAGcYsTyrlLeuSerAsnSerGlnAlaTySerAlaTyTyr 21
Db 42 AATTGCAATTATTAGCTGCTGTACTTGCAGATAATCAAGCTTATCTCATACCAT 101
Qy 22 IleLeuLys 24
            |||||

```

```

Db 102 ATTTTAAAG 110
RESULT 7
LOCUS      BQ138306                      656 bp    mRNA    linear    EST 26-APR-2002
DEFINITION NF001G11PH1086 Phoma-infected Medicago truncatula cDNA clone
ACCESSION  BQ138306
VERSION    BQ138306.1 GI:20274432
KEYWORDS   EST.
SOURCE     Medicago truncatula (barrel medic)
ORGANISM   Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE  1 (bases 1 to 656)
AUTHORS   Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,
            Gonzales,R.A., Bell,C.J., Inman,J.T., Waugh,M.E., Sullivan,J.P.,
            May,G.D. and Paiva,N.L.
TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula Phoma-infected library
JOURNAL    Unpublished
COMMENT    Contact: Paiva NL
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7317
            Fax: 580 221 7380
            Email: nlpaiva@noble.org
            Insert Length: 656 Std Error: 0.00
            Plate: 001 row: G column: 11
            Seq primer: TCACACGAGGAACAGCTATGAC.
FEATURES   source
            1..656
                Location/Qualifiers
                /organism="Medicago truncatula"
                /mol_type="mRNA"
                /db_xref="taxon:3880"
                /clone="NF001G11PH"
                /tissue_type="leaf"
                /dev_stage="Pathogen-induced, young trifoliolate"
                /note="Vector: pBluescript SK(-); Young trifoliolate leaves
                of Medicago truncatula were excised and dip-inoculated in
                a spore suspension of Phoma medicaginis, and incubated in
                humid dishes. Pools of leaves were harvested at 0, 15, and
                30 minutes and 1, 2, 3, 5, 14, 24, 48, 72, and 96, hours,
                and used to prepare total RNA. cDNA was prepared from
                total RNA enriched, pooled samples of equivalent amounts of
                polyA+ enriched, pooled samples of equivalent amounts of
                total RNA from each sample. The cDNA was directionally
                ligated into the Uni-Zap XR vector (Stratagene) and
                packaged using the Gigapack III Gold packaging extracts.
                Phagemids containing cDNA inserts were in vivo excised
                from the recombinant Uni-ZAP XR vector using ExAssist
                helper phage and the E. coli strain Xli-Blue MRF' using SOLR
                cells."
BASE COUNT  177 a 145 c 92 g 242 t
ORIGIN
Alignment Scores:
Pred. No.: 7.67e-05 Length: 656
Score: 99.00 Matches: 19
Percent Similarity: 91.30% Conservatve: 2
Best Local Similarity: 82.61% Mismatches: 2
Query Match: 81.82% Indels: 0
DB: 13 Gaps: 0

US-10-036-492-6 (1-24) x BQ138306 (1-656)

Qy 2 AsnLeuGlnLeuLeuAlaAAGcYsTyrlLeuSerAsnSerGlnAlaTySerAlaTyTyr 21
Db 207 AATTGCAATTATTAGCTGCTGTACTTGCAGATAATCAAGCTTATCTCATACCAT 266
            |||||

```

Qy 22 IleuLeuys 24
Db 267 ATTITAAAG 275

RESULT 8
B78168 622 bp DNA linear GSS 16-JAN-1998
LOCUS T31120TP TAMU Arabidopsis thaliana genomic clone T31120, genomic
DEFINITION survey sequence.

ACCESSION
B78168
VERSION
B78168
SOURCE
GSS.
ORGANISM
Arabidopsis thaliana (thale cress)

REFERENCE
1 (bases 1 to 622)
AUTHORS
Berry, K., Granger, D., Suh, E., Wible, C., Adams, M.D. and Venter
, J.C.

TITLE
A BAC End Sequence Database for Identifying Minimal Overlaps in
Arabidopsis Genomic Sequencing. Update 3

JOURNAL
COMMENT
Other GSSs: T31120TR
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13-21
Class: BAC ends
High quality sequence stop: 622.

FEATURES
source
1..622
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T31120"
/sex="hermaphrodite"
/clone_lib="TAMU"
/note="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII
; Produced by Rod Wing"

BASE COUNT 151 a 107 c 116 g 248 t

ORIGIN
Alignment Scores:
Pred. No.: 0.00114 Length: 622
Score: 91.00 Matches: 18
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 3
Query Match: 28 Indels: 0
DB: Gaps: 0

US-10-036-492-6 (1-24) x B78168 (1-622)

Qy 1 ValAsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyr 20
Db 187 GTTAATTTGCAGCTATTAGCCACCAGCTACCTGCAGCAATAATCAAGCTTACAGTGCATAT 246

Qy 21 TyrIleLeuLys 24
Db 247 CATCTGCTAAAG 258

RESULT 9
B1176998 467 bp mRNA linear EST 07-MAR-2003
LOCUS B1176998
DEFINITION EST17943 cSTE Solanum tuberosum cDNA clone cSTE8011 5' sequence,
mRNA sequence.

ACCESSION
B1176998
VERSION
B1176998
SOURCE
Solanum tuberosum (potato)
ORGANISM
Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 467)
AUTHORS
van der Hoeven, R., Bezzerides, J., Bachem, C., Visser, R., Cho, J.,
Chiemingo, A., Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and
Baker, B.
Generation of ESTs from in vitro grown microtubers
Unpublished
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13F-R.
Location/Qualifiers
1..467
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="cSTE8011"
/tissue_type="axillary buds of stem explants; growing
sink-tubers"
/dev_stage="7, 8 and 10 days"
/lab_host="SOLR"
/clone_lib="cSTE"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue supplied by Christian Bachem and Richard
Visser (Department of Plant Breeding, Wageningen
University, The Netherlands). The cSTA libraries will
attempt to capture the induction and initiation/initial
growth of the tuber in an in vitro system as described in
Bachem et al. (Plant Journal, 1996). Small microtubers
develop from axillary buds attached to stem explants when
placed on a high sucrose medium (10%). Visible
morphological changes occur synchronously at day five in
the axillary buds. The first library, cSTA (1-20) consists
of axillary buds harvested on days 1-3. This targets
those genes involved in induction of the microtubers. The
following libraries, cSTA (21-40) and cSTA (41-60),
capture genes involved in tuber initiation and outgrowth.
This library is noted as P3 in Tanksley lab notebooks."

BASE COUNT 116 a 126 c 89 g 136 t

ORIGIN
Alignment Scores:
Pred. No.: 0.00153 Length: 467
Score: 89.00 Matches: 16
Percent Similarity: 86.96% Conservative: 4
Best Local Similarity: 69.57% Mismatches: 3
Query Match: 73.55% Indels: 0
DB: Gaps: 0

US-10-036-492-6 (1-24) x B1176998 (1-467)

Qy 2 AsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyr 21
Db 167 AATATGCAGCTTTTAGCTGCTGCTACCTGCACACCAACAGGCTTATGCTGATCAT 226

Qy 22 IleuLys 24
Db 227 CTCTCAAG 235

RESULT 10
BG593671

LOCUS	710 bp	mRNA	linear	EST 07-MAR-2003
DEFINITION	EST492349 cSTS Solanum tuberosum cDNA clone cSTS511 5' sequence,			
ACCESSION	BG593671			
VERSION	BG593671.1 GI:13611811			
KEYWORDS	EST.			
ORGANISM	Solanum tuberosum (potato)			
ORGANISM	Solanum tuberosum			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.			
AUTHORS	1 (bases 1 to 710) van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemingo, A., Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.			
TITLE	Generations of ESTs from sprouting potato eyes			
JOURNAL	Unpublished			
COMMENT	Contact: Robin Buell The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA Email: potato-array@tigr.org This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL: http://genome.arizona.edu/orders/ Seq primer: M13P-R.			
FEATURES	Location/Qualifiers			
source	1..710			
	/organism="Solanum tuberosum"			
	/mol_type="mRNA"			
	/cultivar="Kennebec"			
	/db_xref="taxon:4113"			
	/clone="cSTS511"			
	/tissue_type="sprouting eyes from tubers"			
	/dev_stage="12-14 weeks post harvest"			
	/lab_host="SOLR"			
	/clone_lib="cSTS"			
	/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."			
BASE COUNT	207 a	168 c	140 g	195 t
ORIGIN				
Alignment Scores:				
Pred. No.:	0.00274	Length:	710	
Score:	89.00	Matches:	16	
Percent Similarity:	86.96%	Conservative:	4	
Best Local Similarity:	69.57%	Mismatches:	3	
Query Match:	73.55%	Indels:	0	
DB:	10	Gaps:	0	
US-10-036-492-6 (1-24) x BG593671 (1-710)				
QY	2 AsnLeuGlnLeuLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyr 21			
Db	159 AATATGACGCTTTTAGCTGGCTGCTACCTGCAACCAACAGGCTTATGCTGCATATCAT 218			
QY	22 IleLeuLys 24			
Db	219 CTCTCAAG 227			
RESULT 11				
LOCUS	B1178043			
DEFINITION	EST518988 cSTE Solanum tuberosum cDNA clone cSTE12B17 5' sequence, mRNA sequence.			
ACCESSION	B1178043			
VERSION	B1178043.1 GI:14643854			
KEYWORDS	EST.			
SOURCE	Solanum tuberosum (potato)			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			

Tue Feb 3 10:50:18 2004

Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 727)

REFERENCE
AUTHORS
van der Hoeven, R., Berzider, J., Ewing, E., Cho, J., Chiening, A.,
Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.
Generations of ESTs from dormant potato tubers
Unpublished
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13P-R.

FEATURES
source
Location/Qualifiers
1..727
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="GSTDC9"
/tissue_type="dormant tuber"
/dev_stage="one month post-harvest"
/lab_host="SOLR"
/clone_lib="cSTP"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; This library targets genes expressed in dormant
tubers. This library was made from sections of dormant
tuber, avoiding the buds and epidermis. Tubers were stored
for one month post-harvest at 40C. The tuber was peeled,
well away from the surface. Then it was chopped into 1-2
mm cubes and immediately frozen in liquid nitrogen. This
library is noted as P4 in Tanksley lab notebooks."

BASE COUNT 201 a 181 c 143 g 202 t
ORIGIN

Alignment Scores:
Pred. No.: 0.00283 Length: 727
Score: 89.00 Matches: 16
Percent Similarity: 86.96% Conservative: 4
Best Local Similarity: 69.57% Mismatches: 3
Query Match: 73.55% Indels: 0
DB: 12 Gaps: 0

US-10-036-492-6 (1-24) x BG887406 (1-727)

QY 2 AsnLeuGlnLeuAlaAargCystyrLeuSerAnsSergInAlaTySerAlaTyTyr 21
Db 173 AATATGACGCTTTTAGTGGCTGCTACCTGCACACACAGGCTTATGCTGCATATCAT 232

QY 22 IleLeuLys 24
Db 233 CTTCCTCAG 241

RESULT 13
BQ047257
LOCUS
DEFINITION
BQ047257 P. infestans-challenged potato leaf, incompatible
reaction Solanum tuberosum cDNA clone BPL117M13 5' end, mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 742)

Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 742)

REFERENCE
AUTHORS
Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukoianov, A.,
Rangel, P., Haberlach, G.T., Karaycheva, S.A., Tsai, J., Chiening, A.,
Bougri, O., Buell, C.R., Ronning, C.M., Helgeson, J. and Baker, B.
Generation of ESTs from Potato Leaves Challenged with Phytophthora
infestans, Incompatible Interaction (2002)
Unpublished
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.

FEATURES
source
Location/Qualifiers
1..742
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="BPL117M13"
/tissue_type="leaf"
/dev_stage="6 week old"
/lab_host="SOLR"
/clone_lib="P. infestans-challenged potato leaf,
incompatible reaction"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: UC Berkeley, PCEC; sequencing: The
Institute for Genomic Research. Whole plants were
challenged with 450,000 sporangia/ml P. infestans isolate
US-1 (US940501) in Biotron (Madison, Wisconsin). Leaf
tissue was collected at 1, 2, 5, 12, and 24 hours
post-challenge and frozen in liquid nitrogen immediately
upon removal. Kennebec plants showed no signs of HR.
Katahdin plants (susceptible to P. infestans US-1) were
used as controls and showed infection. NOTE: We cannot
exclude the possibility that this sequence is actually
derived from Phytophthora rather than potato."

BASE COUNT 209 a 175 c 152 g 206 t
ORIGIN

Alignment Scores:
Pred. No.: 0.00291 Length: 742
Score: 89.00 Matches: 16
Percent Similarity: 86.96% Conservative: 4
Best Local Similarity: 69.57% Mismatches: 3
Query Match: 73.55% Indels: 0
DB: 12 Gaps: 0

US-10-036-492-6 (1-24) x BQ047257 (1-742)

QY 2 AsnLeuGlnLeuAlaAargCystyrLeuSerAnsSergInAlaTySerAlaTyTyr 21
Db 142 AATATGACGCTTTTAGTGGCTGCTACCTGCACACACAGGCTTATGCTGCATATCAT 201

QY 22 IleLeuLys 24
Db 202 CTTCCTCAG 210

RESULT 14
AG226554/c
LOCUS
DEFINITION
AG226554
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lotus japonicus
Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
1

Lotus japonicus
Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
1

AUTHORS
 Sato, S., Nakamura, Y. and Tabata, S.
TITLE
 Lotus japonicus BAC End sequences
JOURNAL
 Published Only in Database (2002)
REFERENCE
 2 (bases 1 to 639)
AUTHORS
 Sato, S.
TITLE
 Direct Submission
JOURNAL
 Submitted (20-NOV-2002) Shusei Sato, Kazusa DNA Research Institute,
 The First Laboratory for Plant Gene Research; 2-6-7
 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: sato@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/,
 Tel: 81-438-52-3935 (ex.2336), Fax: 81-438-52-3934)

FEATURES
 source
 1..639
 Location/Qualifiers
 /organism="Lotus japonicus"
 /mol_type="genomic DNA"
 /strain="Miyakojima MG-20"
 /db_xref="taxon:34305"
 /clone="LjB14d24.f"
 /clone_lib="genomic BAC library"
 /note="VECTOR: pBelcBAC11"
 BASE COUNT 246 a 134 c 108 g 151 t

Alignment Scores:
 Pred. No.: 0.00946 Length: 639
 Score: 85.00 Matches: 16
 Percent Similarity: 91.30% Conservative: 5
 Best Local Similarity: 68.57% Mismatches: 2
 Query Match: 70.25% Indels: 0
 DB: 29 Gaps: 0

US-10-036-492-6 (1-24) x AG226554 (1-639)
 QY 2 AnLeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyr 21
 DB 208 AATTTCGATTGTAGCTGGCTGTTACTTGGAGTAGAAGCTCATCTGCATACCAT 149

QY 22 IleLeuLys 24
 DB 148 ATCTTAAG 140

RESULT 15
 CB366511 402 bp mRNA linear EST 17-MAR-2003
 ZF001-P00051-DPE-F-A_B08 GISZF001 Danio rerio cDNA clone
 IMAGE:6910684 5', similar to (NM_001256) cell division cycle protein
 27, mRNA sequence.

ACCESSION CB366511 GI:29017162
 VERSION CB366511.1
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
 ; Cyprinidae; Danio.

REFERENCE
 1 (bases 1 to 402)
 Mathavan, S., Wei, C., Thoreau, H., Chia, J.M. and Ruan, Y.
 Genome Institute of Singapore, Zebrafish EST Collection
 Unpublished
 Contact: Ruan Y
 Laboratory of Molecular Biotechnology
 Genome Institute of Singapore
 1 Science Park Road, The Capricorn #05-01, Singapore 117528
 Tel: +65 6827 5200
 Fax: +65 6827 5201

Email: gisry@nus.edu.sg
 GIS Clone ID: ZF001-P00051-PP_C15
 PCR Primers
 FORWARD: M13
 BACKWARD: M13
 Plate: ZF001-P00051-DPE-F-A
 Seq primer: CCGCACTAATGTATACCA
 High quality sequence stop: 402.

FEATURES

source
 1..402
 Location/Qualifiers
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:6910684"
 /issue_type="Embryo"
 /dev_stage="7 Different embryonic stages (From just
 fertilized Embryos to 72 hours just hatched baby fish)"
 /lab_host="PH108"
 /clone_lib="GISZF001"
 /note="Vector: pDNR-LIB; Site 1: Sfi A (GCCATTAGGCC);
 Site 2: Sfi B (GGCGCTCGGCC); Priming method: Sfi-(dT)30
 Primed; Priming sequence: 5-ATTCTAGA GGCGAGCGCGGCC
 GACATG(T)30VN; Directionally cloned, 5' cloning site:
 Sfi A site GCCATTAGGCC; 3' linker/adaptor sequence:
 5-AAGCAGTGGTATCAGCAGAGTGGCC; 3' cloning site: Sfi B
 site GGCGCTCGGCC; 3' linker/adaptor sequence: same
 as the priming sequence; Average insert size: 2Kb; For
 PCR insert analysis: Use M13 Forward and reverse primers;
 Library Amplified Recombinants (inserts): 98%; Library
 complexity: 5x10⁶; Full-length construction (method):
 SMART, a Clontech method; Library constructed by: S.
 Mathavan, Chia-Lin Wei, and Yijun Ruan Genome Institute of
 Singapore"

BASE COUNT 122 a 93 c 96 g 91 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 10.1 Length: 402
 Score: 63.00 Matches: 12
 Percent Similarity: 72.73% Conservative: 4
 Best Local Similarity: 54.55% Mismatches: 6
 Query Match: 52.07% Indels: 0
 DB: 14 Gaps: 0

US-10-036-492-6 (1-24) x CB366511 (1-402)

QY 3 LeuGlnLeuAlaArgCysTyrLeuSerAsnSerGlnAlaTyrSerAlaTyrTyr 22
 DB 51 CTCTCTCTGCTGGCCACGTGCTACTACCGTTACGGGAAGCATACAGGCTACCACTG 110
 QY 23 LeuLys 24
 DB 111 CTGAAA 116

Search completed: February 2, 2004, 17:00:17
 Job time : 4819 secs